City Research Online

City St George's, University of London

Citation: Dimitrova, D. S., Kaishev, V. K. & Zhao, S. (2015). Modeling Finite-Time Failure Probabilities in Risk Analysis Applications. Risk Analysis, 35(10), pp. 1919-1939. doi: 10.1111/risa.12384

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/11951/

Link to published version: https://doi.org/10.1111/risa.12384

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Modelling finite-time failure probabilities in risk analysis applications

Dimitrina S. Dimitrova, Vladimir K. Kaishev*and Shouqi Zhao

Cass Business School, City University London, UK

Abstract

In this paper, we introduce a framework for analyzing the risk of systems failure based on estimating the failure probability. The latter is defined as the probability that a certain risk process, characterizing the operations of a system, reaches a possibly time-dependent critical risk level within a finite-time interval. Under general assumptions, we define two dually connected models for the risk process and derive explicit expressions for the failure probability and also the joint probability of the time of the occurrence of failure and the excess of the risk process over the risk level. We illustrate how these probabilistic models and results can be successfully applied in several important areas of risk analysis among which are systems reliability, inventory management, flood control via dam management, infectious disease spread and financial insolvency. Numerical illustrations are also presented.

Keywords: finite-time failure probability, dependent risk modelling, alarm time, Appell polynomials, dam overtopping

*Faculty of Actuarial Science and Insurance, Cass Business School, City University London, 106 Bunhill Row, EC1Y 8TZ London, UK. Email: v.kaishev@city.ac.uk

1 INTRODUCTION

Research on the probability of systems failure has attracted considerable attention in the risk analysis literature in a variety of contexts, e.g., systems reliability analysis (Todinov 2006; Paté-Cornell and Bea 1992), environmental risk management (Kolen et al. 2013; Serrano-Lombillo et al. 2011, etc.), control of flood risk via dam management (Cox 2009; Kolen et al. 2013), managing inventory and supply chain risk (see e.g. Resurreccion and Santos 2012), managing the risk of insolvency in the financial services industry. A recent contribution in the latter strand of literature is the paper by Gerrard and Tsanakas (2011). Under the assumption of model parameter uncertainty, the authors consider the static problem of setting insurance risk capital at a sufficiently high (threshold) level so that the probability of the total insurance loss exceeding it equals a preassigned small value. In an attempt to describe a more generic situation, the authors interpret this as finding a threshold level which will not be exceeded by a certain risk factor with a predetermined large probability and give two examples from environmental risk analysis, and supply chain risk management where similar situations arise.

While we take inspiration from Gerrard and Tsanakas (2011), in this paper we depart from their focus on parameter uncertainty and make one step further considering a dynamic loss process and time-varying threshold. We address a more general problem of estimating the probability that a risk factor will exceed a certain time-dependent threshold level and its inverse problem of finding a time-dependent threshold level which will not be exceeded by the risk factor with a predetermined large probability. More precisely, we assume that a system is characterized by two stochastic processes. One process models the arrival and severity of some hazardous (critical) events that have negative impact on the system, such as losses caused by system failures or catastrophic collapses. The second one represents the risk mitigation process which involves actions aimed to keep the system operational, such as flood preventive measures, supply of additional inventory to meet demand, capital injections to ensure solvency, etc. It will be convenient for the time being to refer to the two processes as the loss and the gain processes. The latter will be given specific interpretations in section 3, where we demonstrate how the described general model fits particular real life risk analysis problems, e.g. in dam management to mitigate flood risk, in systems reliability risk, in risk of emerging disease spread, and in solvency risk in financial systems. Under this risk framework, we will be concerned with the event that the loss process hits the gain process from below leading to the failure of the whole system,

and we will be particularly interested in the probability that such a failure event occurs within a finite-time interval.

There are several areas of risk analysis where system failure problems arise and can be seen to be covered by the generic risk model introduced above. However, more specifically, depending on the level of stochasticity embedded in the two processes, we will distinguish two versions of the model, which we name the direct and inverse (or dual) risk models, coded as models A and B, respectively. In model A, we assume that systems failures cause losses e.g. running records of flood level, financial liabilities or inventory management costs, which occur at some random instants in time within a fixed time horizon and consider the probability that the aggregate loss resulting from all failures will not exceed the value of a time dependent critical threshold level, modelled by an arbitrary non-decreasing function of time, allowing jumps in the threshold.

Alternatively, we will demonstrate that there are important risk analytic applications where it is relevant to reverse the interpretation of the loss and gain processes in the direct risk model A, formulated in detail in section 2.1. More precisely in the alternative, inverse model B, which is known in the literature as the dual risk model, (see e.g.Asmussen and Albrecher 2010, Dimitrova et al. 2013b), it is assumed that gains (as opposed to losses in model A) which could be e.g., capital gains or inventory replenishment quantities, are of random sizes and arrive at random moments in time, while the cumulative loss (e.g. expense) outgo is modelled by a deterministic threshold function. We consider the probability that the aggregate gain process will go bellow a time dependent critical threshold level, thus causing failure.

The objective of this paper is two-fold. First, we mathematically formulate models A and B in general terms and summarize existing theoretical results obtained under specific assumptions on the probability of system failure within a finite time horizon, which is our primary focus in this paper. A theoretical contribution is then made by deriving a new closed-form expression, given by Proposition B.1, for the joint distribution of the failure time and the deficit level in model A, under the more general assumption of independent, non-identically Erlang distributed loss inter-arrival times (see Appendix B). This expression is new and generalizes previous ruin probability results, obtained by Ignatov and Kaishev (2004, 2012b) for the case of Poisson loss arrivals (see Appendix A). In the second part of the paper, we demonstrate how our proposed models A and B, and the related probabilistic results can be applied as risk analytic tools in various practical risk assessment situations. For the purpose, we illustrate how to interpret the model parameters of our generic models A and B in order

to reflect the specifics of the concrete practical risk assessment problems.

As a main illustrative risk analysis application, in section 3.1, we consider the control of flood risk via managing the risk of dam overtopping. We propose a novel methodology which can aid decision making on multiple water releases, based on the notion of alarm time at which flood warning alarms signal the necessity of water releases. Further examples of risk analysis application of our risk models and results include systems reliability, inventory management, infectious diseases spread and financial insolvency risk.

This paper is organized as follows. In section 2, we introduce the assumptions and notations underlying models A and B in more detail, and provide explicit expressions for the probability that the whole system fails within a finite-time period. We also give an expression for the joint probability of the time of the occurrence of failure and the deficit in model A, a generalization of which is derived in Appendix B. Section 3 is devoted to demonstrating how models A and B and the failure probability results of section 2 can be applied to assess failure risk in systems reliability risk analysis, inventory and supply chain risk management, control of flood risk via dam management, emerging disease spread risk analysis and financial and solvency risk management. Numerical illustrations are also presented in this section. Section 4 provides further discussions on the possible extensions of the applications of model A and B and some concluding remarks.

2 THE RISK MODELLING FRAMEWORK

2.1 Model A

Our generic risk model A is defined as follows. We assume that, as a result of faulty operation of a system, losses, $W_1, W_2, \ldots$, occur at random moments, $T_1, T_2, \ldots$, following an appropriate stochastic counting process, $N(t)$, which counts the number of losses arriving until time t. Different models for the arrival process, $N(t)$:=max$\{i : T_i \leq t\}$, $t > 0$, could be adopted which could capture non-stationarity and clustering in the arrival of losses. As a start, we will assume that $N(t)$ is a homogeneous Poisson process with intensity $\lambda > 0$, i.e. with loss inter-arrival times $\tau_1 = T_1, \tau_2 = T_2 - T_1, \tau_3 = T_3 - T_2, \ldots$, identically, exponentially distributed random variables with parameter λ. Let us note however that a more general setting of non-identical Erlang distributed inter-arrival times is considered in Appendix B.

We will also assume that severities of individual losses are modelled by a sequence of positive

continuous random variables, $W_1, W_2, \ldots$, which are independent of the counting process, $N(t)$. Let Y_i denote their partial sums, i.e. $Y_i = W_1 + \ldots + W_i$, $i = 1, 2, \ldots$, with a joint density function $f(y_1, \ldots, y_i)$. We denote by $S(t)$ the total loss amount up to time t, i.e.

$$S(t) = Y_{N(t)} = \sum_{i=1}^{N(t)} W_i, \tag{1}$$

where we assume that $S(t) = 0$ if $N(t) = 0$. The total loss, $S(t)$, can be viewed as a risk factor whose value is non-decreasing in time. This is a reasonably general model for the total loss and the occurrence and sizes of individual losses from failures. We note that a simpler version of the model, assuming $W_1, W_2, \ldots$ are i.i.d, has been studied in the classical (insurance) risk theory (see e.g. Asmussen and Albrecher 2010). The more general aggregate loss model (1) is better suited for modelling dependence among the consecutive individual losses, which often is the case in real life risk analytic applications as will be illustrated in the next sections.

In what follows, we will be interested in modelling the probability that, within a finite-time interval $[0, x]$, the total loss from the system's failures, $S(t)$, would not exceed the gain process, which is a certain time dependent critical threshold level, modelled by a non-decreasing non-negative deterministic function, $h(t)$ (note that $h(0) \geq 0$ and if $h(0) > 0$, the latter can be interpreted as initial surplus). Such an exceedance event, may be interpreted as the overall failure of the system which may lead to its temporary or permanent inoperability. In fact, it may visually also be interpreted simply as the loss trajectory hitting the critical upper bound, as illustrated in Figure 1. We denote by T, the time when $S(t)$ exceeds $h(t)$ for the first time. Formally one can define the random variable T through the risk process

$$R(t) = h(t) - S(t), \tag{2}$$

as

$$T := \inf\{t : t > 0, R(t) < 0\},$$

and the event of non-failure (overall survival) within the finite-time horizon $[0, x]$ as $\{T > x\}$ (i.e. the random variable T takes value greater than x, including $T = \infty$ when failure never occurs or formally when $R(t) \geq 0$ for all $t > 0$). Alternatively, we will also refer to the event $\{T < x\}$ as the failure event.

Then, we can express the probability that the system does not fail, i.e. survives beyond time x,

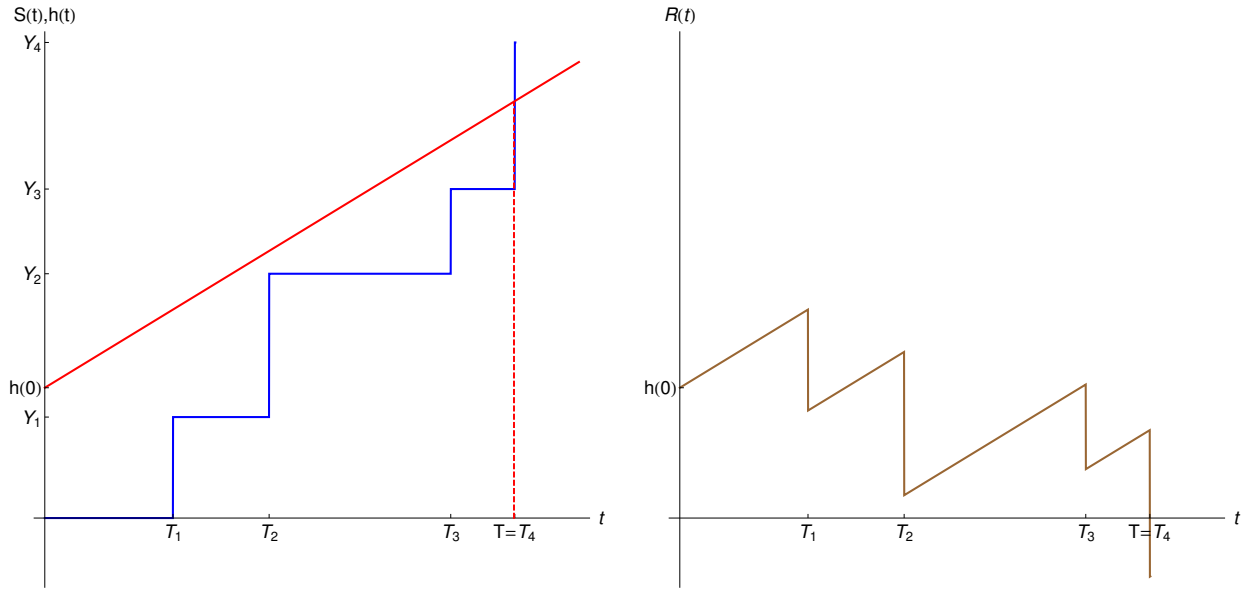

Figure 1: Model A. Left panel: $h(t)$ (red/straight line) represents the deterministic upper critical level and $S(t)$ (blue/staircase line) is the stochastic loss process. Right panel: the risk process $R(t)$.

as $P(T > x)$. In a series of papers, (see Ignatov et al. 2001; Ignatov and Kaishev 2004; Dimitrova et al. 2013a,b) the above mentioned risk model has been studied and the following explicit survival probability formulae have been derived.

$$P(T > x) = e^{-\lambda x}\Big(1 + \sum_{k=1}^{\infty} \lambda^k \int_0^{h(x)} \int_{y_1}^{h(x)} \cdots \int_{y_{k-1}}^{h(x)} A_k(x; \nu_1, \ldots, \nu_k) \times f(y_1, \ldots, y_k) dy_k \ldots dy_1\Big), \quad (3)$$

where $\nu_k = h^{-1}(y_k)$, $h^{-1}(.)$ is the inverse function of $h(t)$, and $A_k(x; \nu_1, \ldots, \nu_k) \equiv A_k(x)$, $k = 1, 2, \ldots$, are the classical Appell polynomials of degree k with a coefficient in front of x^k equal to $1/k!$, defined by

$$\begin{aligned} A_0(x) &= 1, \\ A_k'(x; \nu_1, \ldots, \nu_k) &= A_{k-1}(x; \nu_1, \ldots, \nu_{k-1}), \\ A_k(\nu_k; \nu_1, \ldots, \nu_k) &= 0, \ k = 1, 2, \ldots. \end{aligned} \quad (4)$$

Efficient algorithms for computing formula (3) with a prescribed accuracy have been proposed by Dimitrova et al. (2013a), based on new recurrent expressions for the classical Appell polynomials. It is worth mentioning that a version of formula (3) for the case of integer-valued losses, $W_1, W_2, \ldots$, is

also given in Dimitrova et al. (2013a), where further details on these explicit expressions for $P(T > x)$ and their numerical properties can be found.

It has to be emphasized that, mathematically, the setup in model A is very general. We assume any non-negative non-decreasing gain function $h(t)$ and arbitrarily distributed, possibly dependent, loss severities W_i. We should note however, that the the assumptions on the stochastic nature of the process $S(t)$ and the deterministic function $h(t)$ are specified at the initial time, $t = 0$. Both $S(t)$ and $h(t)$ in practice would be estimated at time $t = 0$, based on historic data on the arrival times T_i, the amounts W_i, $i = 1, 2, \ldots$ and the components of the deterministic cumulative gain $h(t)$. Thus, risk related decisions made at time zero would be static since they would not take into account the future realizations of $S(t)$ and $h(t)$. At the same time, decision making can be made dynamic by splitting the time horizon, x, into shorter sub-periods and taking into account all arriving information on the underlying processes. This approach is illustrated in our main example on flood risk management via water releases according to flood warnings at multiple alarm times (see section 3.1).

Furthermore, the assumption of Poisson loss arrivals, which seems more restrictive than the other assumptions, can also be further generalized to a non-homogeneous Erlang arrival process (see Ignatov and Kaishev 2012b; note Erlang distribution is a special case of Gamma distribution and so, generalizes the exponential inter-arrival times arising from a Poisson process) and to the case where the inter-arrival times follow a linear combination of non-identical exponential distributions (see Dimitrova et al. 2013b).

As mentioned in section 1, in some real life risk situations, the occurrence of the exceedance (failure) event $\{T < x\}$ may not mean an immediate system's failure. Therefore, given failure at time T, we will also be interested in the amount Y, by which the total loss $S(T)$ exceeds the critical risk level $h(T)$, which we call deficit, i.e. in the random variable $Y = S(T) - h(T) \equiv Y_{N(T)} - h(T)$. More precisely, we will assume that the system is still operational if an exceedance event occurs and simultaneously, the deficit Y is below a certain pre-determined (controllable) threshold level $y \geq 0$; and assume that the system fails immediately at time $T < x$ if $Y > y$. Thus, in this setting, we are concerned with the more realistic failure event $\{T < x, Y > y\}$, determined by the joint distribution of the random variables T and Y, and more precisely, with the probability of its occurrence, $P(T < x, Y > y)$. Under the assumptions of Poisson loss arrivals and the setup introduced above, the following explicit expression for $P(T < x, Y > y)$ is obtained by Ignatov and

Kaishev (2012a),

$$P(T<x,Y>y)=\int_y^{+\infty}f(y_1)dy_1-\int_y^{h(x)+y}e^{-\lambda h^{-1}(y_1-y)}f(y_1)dy_1-e^{-\lambda x}\int_{h(x)+y}^{+\infty}f(y_1)dy_1$$
$$+\sum_{k=2}^{\infty}\int\dots\int_{C_k}\left\{B_{k-2}\left(h^{-1}\left(y_{k-1}\right);\nu_1,\dots,\nu_{k-2}\right)-B_{k-1}\left(h^{-1}\left(y_k-y\right);\nu_1,\dots,\nu_{k-1}\right)\right\}f(y_1,\dots,y_k)dy_k\dots dy_1$$
$$+\sum_{k=2}^{\infty}\int\dots\int_{D_k}\left\{B_{k-2}\left(h^{-1}\left(y_{k-1}\right);\nu_1,\dots,\nu_{k-2}\right)-B_{k-1}\left(x;y_1,\dots,y_{k-1}\right)\right\}f(y_1,\dots,y_k)dy_k\dots dy_1, \quad (5)$$

where $C_k=\{(y_1,\dots,y_k):0<y_1<\dots<y_{k-1}\le y_{k-1}+y<y_k<h(x)+y\}$,
$D_k=\{(y_1,\dots,y_k):0<y_1<\dots<y_{k-1}<h(x)\le h(x)+y\le y_k<+\infty\}$,
$B_k(z;\nu_1,\dots,\nu_k)=e^{-\lambda z}\left[A_0+\lambda A_1\left(z;\nu_1\right)+\dots+\lambda^k A_k\left(z;\nu_1,\dots,\nu_k\right)\right]$, and $A_k(z)$ are as in (4) the classical Appell polynomials.

Let us note that, when $y=0$, formula (5) coincides with formula (3) (see Appendix A for a detailed proof). The assumption of Poisson loss arrivals can again be generalized to non-homogeneous Erlang arrivals, which can be adapted to fit an arbitrary distribution via randomization, and an explicit expression for $P(T<x,Y>y)$ in this case is derived in Appendix B.

2.2 Model B

In this section, we provide a detailed description of model B. To distinguish it from model A, we refer to model B as the dual risk model and use the subscript "dual" in the notations. We assume a system starts to operate at an initial surplus level u, $u>0$. Its further operation generates gains and losses. The losses are non-stochastic and arrive continuously following a deterministic function of time, $g(t)$, with the initial condition, $g(0)=0$. The operability of the system is maintained dynamically by a gain process $S(t)$, consisting of a series of consecutive gains, denoted by the random variables $W_1,W_2,\dots$, occurring at random moments, $T_1,T_2,\dots$. The number of gains arriving before time t follows a counting process $N(t)$:=max$\{i:T_i\le t\}$, $t>0$, which is independent of the gain sizes. Again, as a start, we assume that $N(t)$ is a Poisson process with a constant intensity $\lambda>0$, i.e. the inter-arrival times $\tau_1=T_1,\tau_2=T_2-T_1,\tau_3=T_3-T_2,\dots$, are i.i.d. exponentially distributed with parameter λ. Thus, we have

$$S(t)=\sum_{i=1}^{N(t)}W_i,$$

and the risk process is therefore

$$R(t) = u - g(t) + S(t).$$

Note that we can write $R(t) = -h_{\text{dual}}(t) + S(t)$, where $h_{\text{dual}}(t) = -u + g(t)$ and that in contrast to model A, where the initial surplus could be zero, here $u > 0$. Again, we are interested in the probability of the risk process reaching the horizontal axis within a finite period $[0, x]$, representing the loss process hitting the gain process from below (as illustrated in Figure 2) causing the overall failure of the system at time T_{dual}. More precisely, we would be interested in the probability of non-failure, $P(T_{\text{dual}} > x)$, where the instant of the failure (exceedance) event, T_{dual}, is defined through the risk process as

$$T_{\text{dual}} := \inf\{t : t > 0, R(t) < 0\}.$$

Clearly, the definition of T_{dual} can be re-written as

$$\begin{aligned} T_{\text{dual}} &= \inf\{t : t > 0, u - g(t) + \sum_{i=0}^{N(t)} W_i < 0\} \\ &= \inf\{t : t > 0, \sum_{i=0}^{N(t)} W_i < -u + g(t)\} \\ &= \inf\{t : t > 0, \sum_{i=0}^{N(t)} W_i < h_{\text{dual}}(t)\}, \end{aligned}$$

where $h_{\text{dual}}(t) = -u+g(t)$. Thus, the failure event, $\{T_{\text{dual}} < x\}$, is essentially equivalent to a trajectory $S(t)$ hitting a lower risk bound $h_{\text{dual}}(t)$, as illustrated in Figure 2.

As shown by Dimitrova et al. (2013b), an enlightening connection between models A and B can be made by reflecting the left panel of Figure 2 with respect to bisector $y = t$, as demonstrated in Figure 3. Thus, the problem of finding the probability of a trajectory hitting a lower bound is translated to that of finding the probability of a trajectory hitting an upper bound with reverse interpretations of the model parameters in A and B and some minor adjustments (see Lemma 2.1 in Dimitrova et al. 2013b for more details). With such a duality link, the results on the probability of the overall system failure within a finite horizon in model B can be easily translated to and restated in terms of model A.

It is worth noting that under model B, as shown in the right panel of Figure 2, the deficit at the instant T_{dual}, of the risk process crossing the horizontal axis, is zero, i.e. $Y = S(T_{\text{dual}}) - h_{\text{dual}}(T_{\text{dual}}) = 0$.

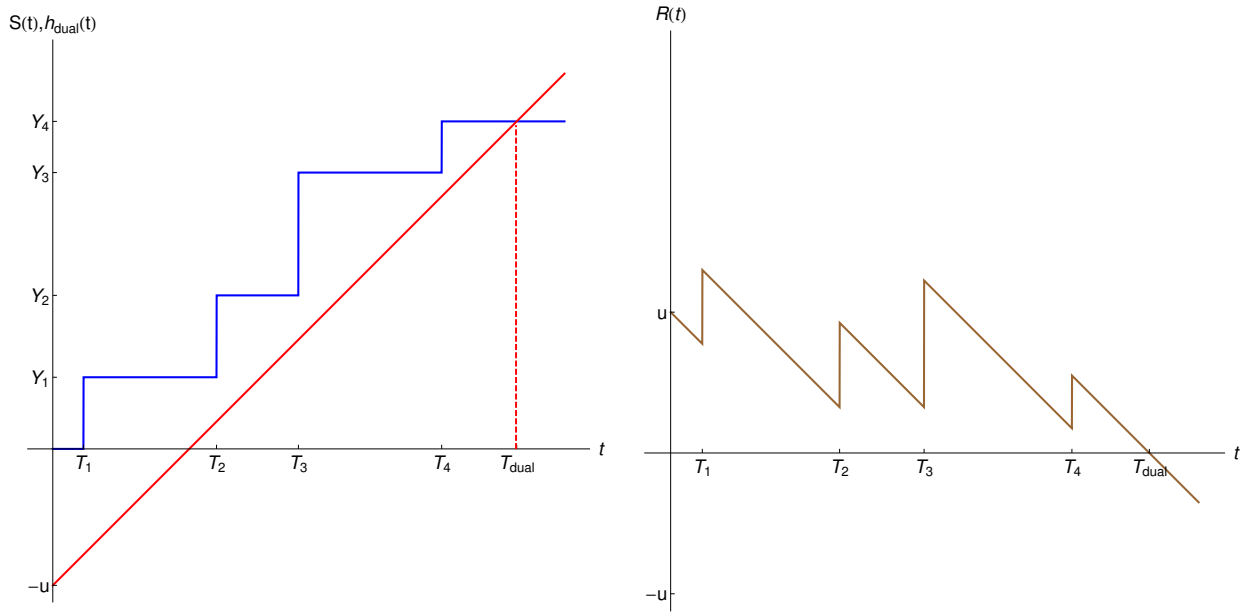

Figure 2: Model B. Left panel: $h_{\text{dual}}(t)$ (red/straight line) represents the lower risk level bound and $S(t)$ (blue/staircase line) is the stochastic gain process. Right panel: the risk process $R(t)$.

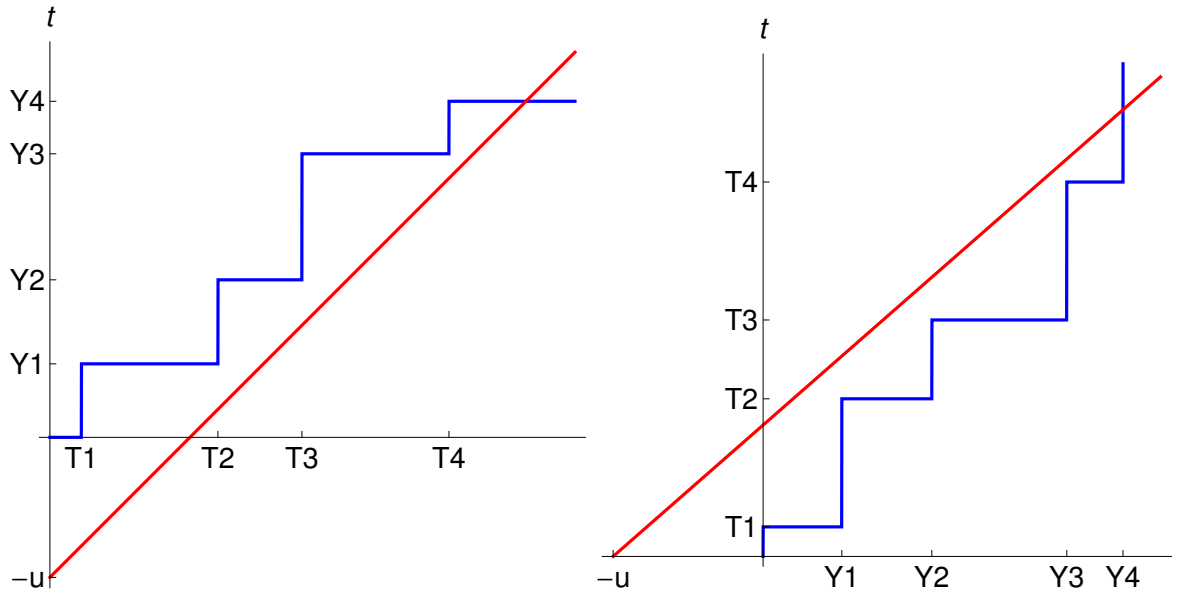

Figure 3: A connection between model A and model B through a reflection along the 45° line.

Therefore, in model B, as opposed to model A, we will not be interested in the deficit, Y, at the failure time T_{dual}.

3 RISK ANALYSIS APPLICATIONS

In this section, we demonstrate that models A and B can be effectively applied as risk analytic tools in analyzing system reliability risk, inventory and supply chain risk management, control of flood risk via dam management, risk of emerging disease spread losses and financial and solvency risk management. For this purpose, without loss of generality, for numerical illustrations we have specified the failure event arrivals to follow a Poisson process and the severities $W_1, W_2, \ldots$ to be independent, identically distributed random variables. For alternative non-Poisson failure event arrivals, e.g. non-identical Erlang inter-arrival times one can implement formula (12) for $P(T < x, Y > y)$ without substantial additional computational cost, as has been demonstrated for instance in Dimitrova et al. (2013b) for the special case of computing $P(T > x)$. Dependence between the consecutive losses, e.g. using copula functions, can also be numerically implemented, as shown in Dimitrova et al. (2013a).

3.1 Control of flood risk via dam management

In this application, we consider the control of flood risk via dam management. Risk analysis in dam management has been intensively explored in the existing literature, see e.g. Lave and Balvanyos (1998), Valdes and Marco (1995), and Cox (2009), to mention only a few of the works. As discussed by Lave and Balvanyos (1998), there are various reasons that may cause dam failures. Here, we focus on the dam failure due to water level overtopping its crest. As has been reported by the US Association of State Dam Safety Officials, failure of dams due to overtopping accounts for approximately 34% of all dam failures in US. Quantitative and probabilistic risk models have been utilized in previous literature as risk analytic tools with a two-fold purpose: for modelling overtopping events, and for decision-making on managing water releases for dams and reservoirs, see e.g. Goodarzi et al. (2011), Serrano-Lombillo et al. (2011), Tingsanchali and Chinnarasri (2001) and Kwon and Moon (2006). Our approach to the first problem is to model the occurrence of an overtopping event using model A, introduced in section 2.1. The latter assumes a stochastic process modelling the occurrence of record high water levels, and a critical threshold level close below the crest of the dam. We introduce the critical level, which if reached by the water signals a warning alarm. This is a core part of our

methodology, proposed as a solution to the second problem of managing water releases. It allows for dynamically determining appropriate instants for water releases to efficiently mitigate the risk of overtopping. In order to implement this novel methodology, as we show, it is essential to be able to evaluate the probability that, within a fixed time period, the running record water level will exceed the critical threshold, and also the probability of overtopping the crest within such a period. The latter is shown to be expressed as the probability of the running record water level exceeding the critical threshold by an amount greater than the difference between the critical threshold level and the crest. We demonstrate that it is possible to evaluate these probabilities using the probabilistic results presented in Section 2. Following is a more formal introduction of our flood risk mitigation methodology.

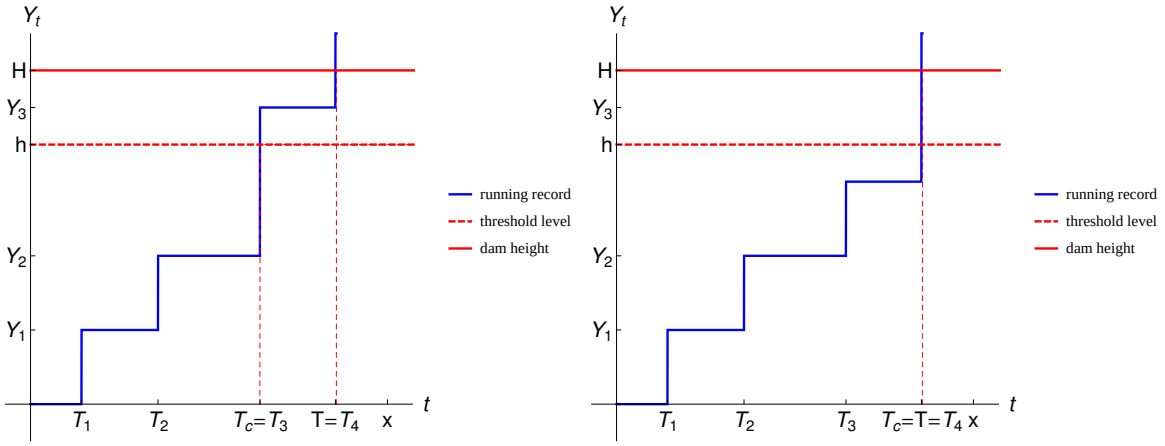

Figure 4: A stochastic model for the risk of overtopping.

It often happens that, as a result of heavy rains, flood water levels in dams increase rapidly, almost jump-wise. This suggests that such a process can be modelled by a pure jump stochastic process, $S(t)$, defined as in (1), where flood jumps in the water level are of unknown random sizes, $W_1, W_2, \ldots$, occurring at some random moments $T_1, T_2, \ldots$ (see Figure 4). Let us note that it is relevant to assume only positive (upward) jumps, i.e. $W_i > 0$, since negative (downward) jumps due to water releases for various other purposes (e.g. irrigation) are usually deterministic and can be considered negligible compared to huge inflows of flood water in the dams, causing upward jumps in the water level. The record levels of such a process can be denoted as $Y_1 = W_1, Y_2 = W_1 + W_2, \ldots$ (see Figure 4). It is then natural to assume that there is an underlying process $N(t) :=\max\{i : T_i \leq t\}$, $t > 0$, counting the occurrences, $T_1, T_2 \ldots$ of the running records, $Y_1, Y_2, \ldots$, before time t. In accordance with Cheng et al. (1993), we assume that $N(t)$ follows a homogeneous Poisson process with intensity

λ, although our modelling framework A (see section 2.1) allows other processes to be considered when calibrating the model in practice. The event of overtopping within a time period $[0, x]$ would then occur at time $T < x$, when the running record water level, $Y_{N(T)}$, exceeds the crest of the dam, denoted by H (see Figure 4). This corresponds to the definition of T through the risk process, $R(t)$, introduced in (2) (see section 2). As an overtopping risk preventive measure, it is natural to consider a critical (warning) threshold water level $h(t)$, lower than the crest, H, to signal the necessity of water release in the event of the exceedance $Y_{N(t)} > h(t)$. If we denote by T_c the instant at which the exceedance event, $Y_{N(T_c)} > h(T_c)$, occurs, we can then evaluate its probability, $(1 - P(T_c > x))$, using formula (3) of section 2 with T replaced by T_c. In general, our model allows the warning level, $h(t)$ to vary over time, reflecting particular environmental circumstances, e.g. current weather and precipitation levels and their forecasts, but without loss of generality here we will assume a constant level, $h(t) = h < H$. As illustrated in the left and right panels of Figure 4, overtopping may result following two alternative scenarios. However, trajectories such as that in the left panel are not of interest, as water can be directly released when the running record, $Y_{N(T_c)}$, exceeds the threshold level h at time T_c, to avoid overtopping. Thus, we only consider trajectories that cause overtopping, as illustrated in the right panel of Figure 4, which corresponds to the exceedance, $Y_{N(T_c)} - h$, being greater than $H - h$. Therefore, it is also important to embed the exceedance, $Y = Y_{N(T_c)} - h$, in our modelling framework. We note that the exceedance, $Y = Y_{N(T_c)} - h$, directly corresponds to what we introduced earlier in section 2.1 as the deficit Y. It is now natural to consider the probability that an overtopping at time $T \equiv T_c$ will occur within a finite time period $[0, x]$, which is equivalent to considering the probability, $P(T_c < x, Y > y)$, that the exceedance event, $Y_{N(T_c)} > h$, occurs, and the exceedance, $Y = Y_{N(T_c)} - h$, is greater than $y = H - h$. The latter overtopping probability, $P(T_c < x, Y > y)$, can be computed using formula (5) of section 2.

In what follows, we first numerically illustrate the evaluation of the probability, $P(T_c < x)$, of exceeding the critical level h, and the overtopping probability $P(T_c < x, Y > H - h)$ (see Example 3.1). This allows us to comment on the sensitivity of the latter with respect to the choice of the critical level h. Then, we concentrate on demonstrating how the two probabilities naturally arise in the definition of a flood warning alarm time, which is at the core of our proposed methodology for flood water release.

Example 3.1 (Control of flood risk via dam management.)

For the purpose of this example, we consider two alternative scenarios, one with small/medium-sized flood jumps in the water level, $W_1, W_2, \ldots$, occurring more frequently, and a second scenario with possibly large jumps, $W_1, W_2, \ldots$, occurring with a lower frequency. In the first scenario, the jump sizes are assumed independent, exponentially distributed with parameter α, i.e. $W \sim$Exp(α), and occur according to a Poisson process with a high intensity λ. In the second case, λ is assumed low and in order to model high-severity jumps, we assume the W_i's are i.i.d. following a generic Pareto distribution with shape parameter θ and scale parameter β, with probability density function

$$f_W(z) = \frac{\theta \beta^\theta}{(z+\beta)^{\theta+1}}, \quad z > 0,$$

which belongs to the family of heavy tailed distributions. With a selected set of parameter values, we compute the probabilities $P(T_c < x)$ and $P(T_c < x, Y > H - h)$ for varying x, see Figure 5. In order to make a fair comparison, we have equated $\lambda\mu$ in the two situations, where μ is the mean of W_i, so that the expected total water inflow severity per unit of time is identical in the two scenarios. As we assume small μ in the first scenario, it is not very likely to have a single water inflow that causes an exceedance event with severity $Y = Y_{N(T_c)} - h(T_c) > H - h$, and so $P(T_c < x, Y > H - h)$ is close to zero. Thus, we compute and plot the overtopping probability, $P(T_c < x, Y > H - h)$, only for the case of Pareto distributed W's.

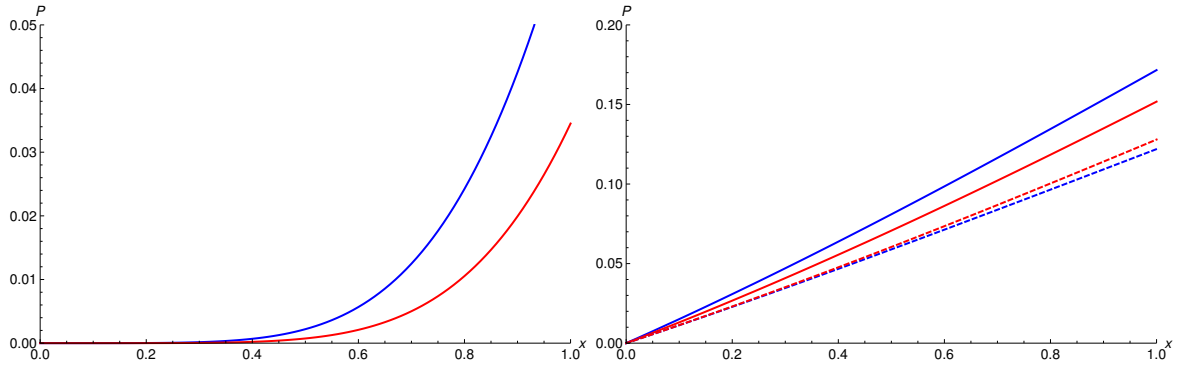

Figure 5: The probabilities $P(T_c < x)$ (solid lines) and $P(T_c < x, Y > H - h)$ (dashed lines) against varying x. Left panel: exponential water inflow severities with parameter values $\lambda = 10$ and $\alpha = 0.2$. Right panel: Pareto water inflow severities with parameter values $\lambda = 1$, $\beta = 50$ and $\theta = 2$. Other parameter values: $H = 100$, $h_1 = 0.8H$ (blue lines) and $h_2 = 0.9H$ (red lines).

As noted, the expected total water inflow severity per unit of time is the same but because the Pareto distribution is heavy-tailed, the exceedance probability, $P(T_c < x)$ for any fixed value of x, is clearly higher in this case (see Figure 5). Furthermore, the exceedance probability $P(T_c < x)$ is higher

for any value x in the case of $h_1 = 0.8H$, which is natural to expect since $h_1 = 0.8H < h_2 = 0.9H$. As opposed to the exceedance probability, the overtopping probability $P(T_c < x, Y > H - h)$ is lower for $h_1 = 0.8H$, which is also to be expected, because a lower h leads to a higher $H - h$ and therefore the event $Y > H - h$ of overtopping is less likely, see the right panel of Figure 5. This illustrates the importance of the choice of the critical level h and its impact on the overtopping probability. We have selected two different values of h to illustrate that a low critical level h_1 leads to a higher probability of sounding an alarm but such a warning is less "alarming" as $H - h_1$ is larger which leaves more space before the crest of the dam is reached. As demonstrated, a high threshold level h_2 can be more indicative of the risk of overtopping but would leave insufficient time for the projected water release to take place before an overtopping event occurs, which can be very risky and therefore is not desirable. Clearly, a balanced choice of the threshold level h should be sought.

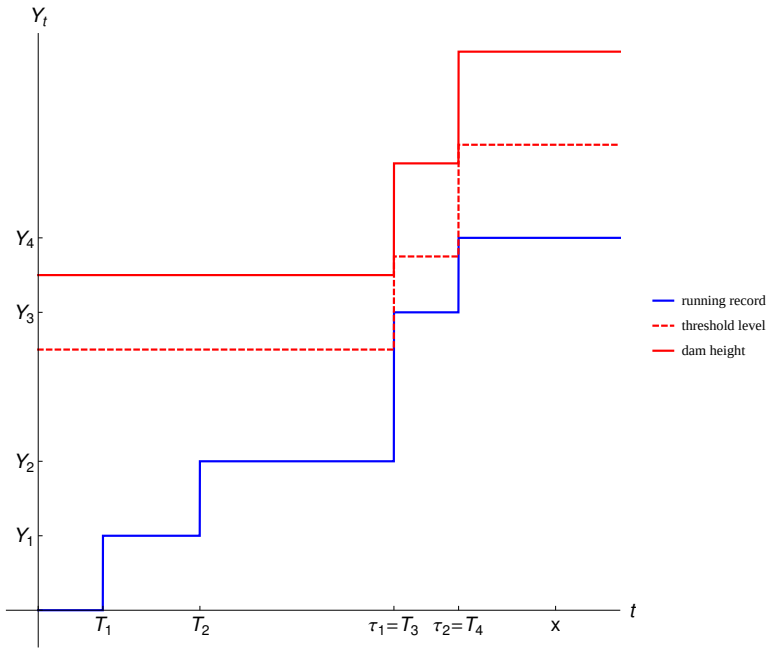

Figure 6: A direct strategy of dam management where water is released instantaneously at the time instants when the pre-specified critical threshold h is exceeded.

Next, we focus our attention on the second problem of managing water releases. Here, we consider a series of water releases of deterministic amounts at certain moments, $\tau_1, \tau_2, \ldots$, within a finite-time horizon $[0, x]$, in order to avoid overtopping events and prevent loss of lives and damage to property as a result of overtopping. In order to incorporate water releases in our model, we assume that the real time needed for a water release is negligibly small, i.e. an appropriate (deterministic) amount of water is released instantaneously, and thus preventing overtopping. Mathematically, such

deterministic downward (negative) jumps in the water level process are equivalent to upward jumps of the same sizes in the upper bounds representing the threshold level $h(t)$ and the height of the dam H, as illustrated in Figure 6. Of course this is only an equivalent mathematical transformation of the model. In reality, the height of the dam H is fixed and cannot be increased. However, for the purpose of computing the related overtopping probability, such a formal mathematical transformation is very helpful.

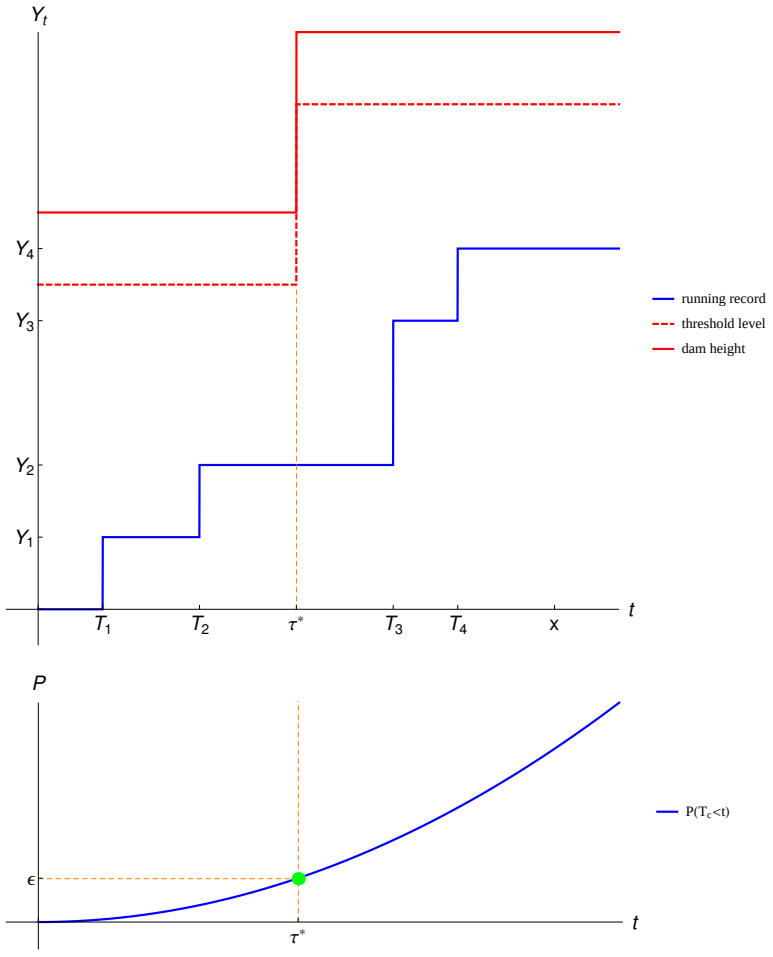

Figure 7: A probabilistic strategy of dam management incorporating an early warning signalling for water releases based on the exceedance probability $P(T_c < t)$.

A natural question then arises. "When is the appropriate time to release water?" A direct answer is not to release water until the water reaches the threshold level h at time T_c, as illustrated in Figure 6. As mentioned earlier, the event of the water level exceeding the pre-specified critical threshold h could serve as an early warning that the water level is alarmingly high and water should be released

to prevent overtopping. However, this may not be a wise idea since the time T_c of water release is a random variable and therefore is not predictable at time $t = 0$. In addition, it may be argued that it is too late and risky to release water at the exceedance time T_c, which will not reduce the likelihood of the occurrence of an overtopping event at time T_c. More precisely, this warning strategy only works if the water level exceeds the critical threshold level h but does not reach the crest of the dam H, i.e. as illustrated by the scenario in the left panel of Figure 4. Should the overtopping event occur at time T_c (see the right panel of Figure 4), this strategy does not suggest any warning or preventive measures prior to the overtopping event. It is therefore not a very efficient way of mitigating overtopping risk.

An alternative strategy that could be pursued is to determine the moment for a water release based on the value of $P(T_c < t)$, as illustrated in Figure 7. Clearly, $P(T_c < t)$ is a monotonically increasing function of t with $P(T_c < 0) = 0$. Thus, it is possible to find the instant τ, such that the probability, $P(T_c < \tau)$, of the water level hitting the threshold level h within a finite period $[0, \tau]$, reaches a prescribed (critical) value ϵ, $0 < \epsilon \ll 1$, when the likelihood of an exceedance event, $Y_{N(t)} > h(t)$, becomes significant. This is the forecasted time at which water should be released. Depending on the parameter settings and the value of ϵ, τ can sometimes be larger than x. In such a case, water should be released at time x. Mathematically, such an alarm time is defined as

$$\tau = \sup \{0 < t \leq x : P(T_c < t) \leq \epsilon\} . \tag{6}$$

In contrast to the direct strategy of releasing water at the (random) time T_c, here τ is not random and is predetermined at time $t = 0$. As can be seen from (6), in order to compute τ, it sufficies to be able to compute the exceedance probability $P(T_c < t)$, which as demonstrated in Example 3.1 can be done using formula (3) of section 2.

With definition (6), we revisit Example 3.1, and compute the alarm times τ with varying threshold level h, as plotted in Figure 8, assuming the flood jumps W_i are independent, Pareto distributed. As can be seen, the alarm time depends on the prescribed (critical) probability ϵ and on the choice of h, which reemphasizes the remark we have made earlier that the pre-specified threshold level h should be selected carefully in order to be used effectively in providing early warning signals. It is worth noting that, because we require two correct digits after the decimal point when computing the alarm time τ, the curves appear like a staircase. They should be smooth should an infinite accuracy of τ be pursued.

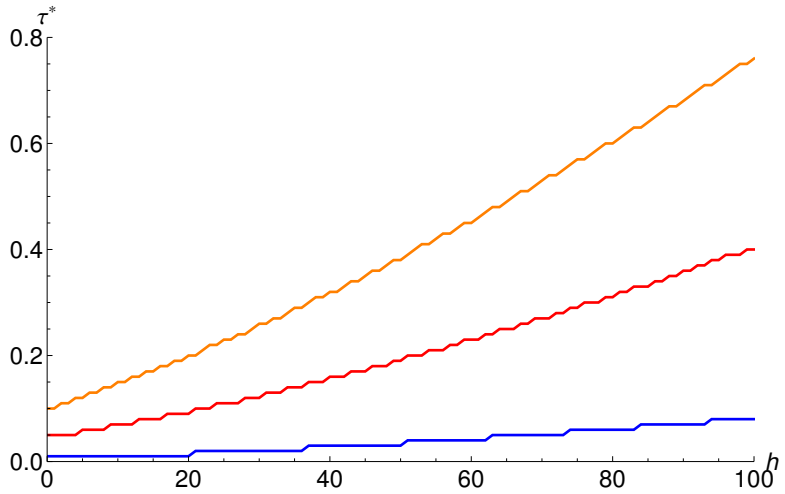

Figure 8: Alarm times against varying threshold levels. Blue, red and orange lines represent $\epsilon = 0.01, 0.05, 0.1$ respectively. Parameter values: $\lambda = 1$, $\beta = 50$, $\theta = 2$ and $H = 100$.

Following the definition of a single alarm (6), it is also possible to define a system with multiple alarms. Suppose the first alarm is sounded at time τ_1 to warn the water level is alarmingly high. Reasonably, the projected water release should take place to decrease the likelihood of an overtopping event. Suppose at each alarm time $\tau_i < x$, $i \geq 1$, the water height is lowered by some deterministic amount v_i, which as we explained earlier can be viewed equivalently as an instantaneous upward jump of the same size v_i in the upper bound. Thus, the upper bound $h(t)$ becomes a process with pure jumps of known sizes v_i at deterministic moments τ_i, starting at $h(0) = h$. Given survival until time τ_i, the next alarm time τ_{i+1} can be defined as

$$\tau_{i+1} = \sup\left\{\tau_i < t \leq x : P(T_c < t | T_c > \tau_i) \leq \epsilon_{i+1}\right\}, \quad i \geq 1, \tag{7}$$

where $\{\epsilon_i\}_{i\geq 1}$ is a sequence of prescribed critical probability levels, $\epsilon_i > 0$, $i = 1, 2, \ldots$. For simplicity, one may use the same value of ϵ for consecutive alarm times, i.e. $\epsilon_i = \epsilon$.

It has to be noted that the second strategy proposed here to release water at time τ_i, based on definition (7), also does not provide a perfect prevention from an overtopping event. Clearly, this strategy is probabilistic, and it does not completely prevent an overtopping event occurring before x, although the overtopping probability is very small (smaller than ϵ). Certainly, smaller ϵ might be used to further diminish the overtopping probability. However, this would lead to a higher frequency of water release, which is again not desirable.

A more practical way of dam management could be to employ a hybrid strategy of the two

alternatives discussed above. More precisely, a series of alarm times τ_i for water releases can be computed at the start time $t = 0$ based on (7), and the actual water level would also need to be inspected over time. Thus, water should be released at each alarm time τ_i to prevent overtopping, and in the very unlikely event of the prescribed threshold level h being exceeded, water would be release immediately.

Finally, another remark should be made that the idea of employing an early warning system in risk analysis has been extensively explored in previous literature, see e.g. Paté-Cornell (1986), Das and Kratz (2012), etc. It is worth noting that the way of defining the early warning (alarm) times is not unique. Here, we have considered the finite-time exceedance probability, $P(T_c < t)$, as the central quantity determining the alarm time. Other quantities, such as the size of a single water inflow severity (Probable Maximum Flood, see e.g. Bullard (1986)), or the joint probability $P(T_c < t, Y > H - h)$, can also be considered when devising alarm time systems.

3.2 Other risk analysis applications

3.2.1 Systems reliability risk

In this application, losses, $W_1, W_2, \ldots$, occur at random moments in time, $T_1, T_2, \ldots$ and may vary in size, from relatively small but multiply occurring losses, e.g. slowdowns and server failures in computer networks disrupting electronic business transactions, to possibly huge losses due to not so frequent but severe failures. In general, as noted by Todinov (2006), losses can be measured in different units, e.g. in lost production time or volumes of defective products, lost lives, financial losses due to credit defaults, insurance claims or tax avoidance, amount of hazardous chemicals released in the environment, warranty payments and lost customers, etc. As has been established by Todinov (2006) "maximizing the reliability of a system, does not necessarily minimize the losses from failures". The author considers an interesting alternative approach to reliability allocation based on maximizing the profit expected to emerge from the operation of the system. It can be argued that this problem, and the related problem of assuring that profit stays positive within a fixed time interval with a very high probability, can be addressed applying the risk theoretic approach, formalized as model A in section 2.1. Thus, the problem of interest, which can be viewed as an extension of the model and profit optimization problem considered by Todinov (2006), can be stated as follows.

Problem 1. Find an appropriate "aggregate gain" function $h(t) \in \mathcal{H}$, such that the probability

of the system's non-failure (survival)

$$P(T > x) = 1 - P\left(\inf_{0 \le t \le x} (h(t) - S(t)) < 0\right) \ge 1 - \epsilon, \tag{8}$$

where $\epsilon > 0$ is a sufficiently small preassigned value, $\mathcal{H}$ is a rather general class of non-decreasing possibly discontinuous functions, $h(t)$, which may have jumps (e.g. a practically important case is step-wise constant $h(t)$), and the loss severities W_i, $i = 1, 2, \ldots$, can be integer-valued or continuous, with any joint loss distribution, including multivariate dependent distributions, e.g. any copula construction.

We note that, depending on the choices of $\mathcal{H}$ and ϵ, the solution for Problem 1 may not exist, or if it exists, it may not be unique. Therefore, one may sometimes pursue an optimal solution in some sense, e.g. maximizing the finite-time survival probability in (8). It is also worth mentioning that, because $\mathcal{H}$ is a fairly general class of functions, it may not be possible to solve Problem 1 analytically and numerical approaches may need to be considered. In order to solve Problem 1 numerically, formulas (3) and (5) can successfully be applied. We illustrate how this is done in the next example, by imposing restrictive assumptions on the class $\mathcal{H}$.

Example 3.2 (Systems reliability risk.)

In this example, we assume the loss severities W_i, $i = 1, 2, \ldots$, follow an i.i.d. logarithmic distribution with parameter α, i.e. $W \sim \text{Log}(\alpha)$, with generic probability mass function $P(W = i) = -\alpha^i/(i \ln(1-\alpha))$. $N(t)$ is assumed a homogeneous Poisson process with constant intensity λ. We denote by u, $u > 0$, the initial surplus level of the system, i.e. $h(0) = u$, and consider $\mathcal{H}$ as a class of piecewise functions with an upward jump of fixed size J at time t_J, $t_J \in [0, x]$, from which we select $h(t)$, which is the solution to Problem 1. Including a jump of size J in the gain function $h(t)$, which may occur at a variable time t_J, can be interpreted as a lump-sum capital injection to avoid the system failure within the time period $[0, x]$. We assume that the critical risk level $h(t)$ has different slopes before and after the jump time t_J, denoted respectively by c_1 and c_2, where c_1 is fixed and c_2 is determined by the instant of the jump t_J. This is to ensure that all gain functions $h(t)$ belonging to $\mathcal{H}$ reach the

same gain level at the end of the time period, x. Mathematically, $\forall h(t) \in \mathcal{H}$, we have

$$h(t) = \begin{cases} u + c_1 t, & 0 \leq t < t_J, \\ u + c_1 t_J + J + c_2(t - t_J), & t_J \leq t \leq x. \end{cases} \tag{9}$$

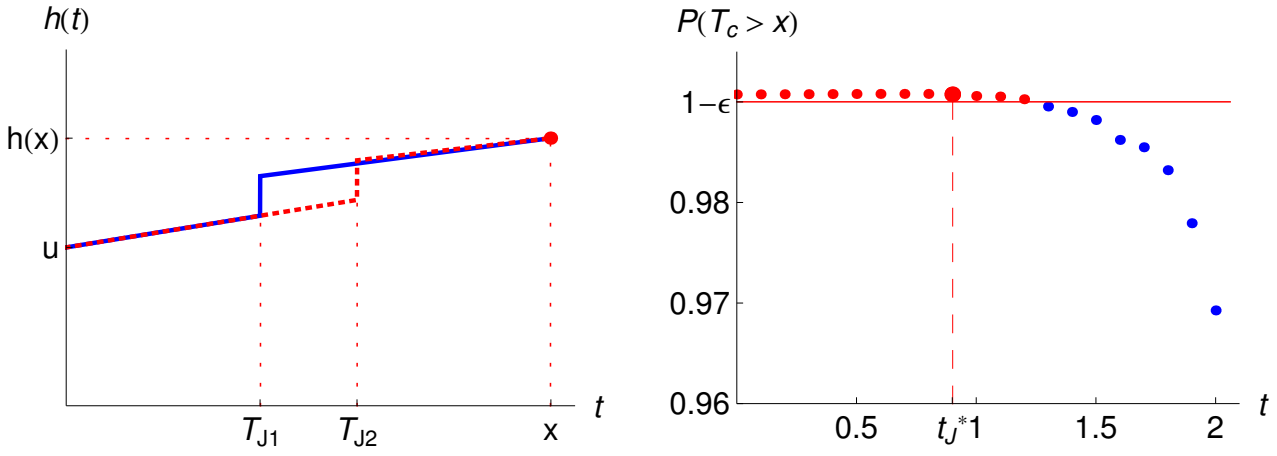

Figure 9: Left panel: two possible candidates $h_1(t), h_2(t) \in \mathcal{H}$. Right panel: $P(T > x)$ as a function of the location t_J of the jump of size J. Parameter values: $\lambda = 5$, $\alpha = 0.1$, $u = 12.5$, $c_1 = 3$, $x = 2$, $J = 2.5$ and $\epsilon = 0.01$.

In the left panel of Figure 9, we plot two possible candidates from the class $\mathcal{H}$ for illustrative purposes. As discussed, both choices of the gain functions $h(t)$ reach the same level of $h(x)$ at the end, x, of the time interval considered, and the after-jump slope varies according to the location of the jump. For a selected parameter set, we evaluate the finite-time survival probability with the discrete version of formula (3), utilizing the fact that formula (3) is valid for any non-negative non-decreasing function allowing jumps, and numerical results are summarized in the right panel of Figure 9. As can be observed, there are many functions $h(t)$ from the chosen class $\mathcal{H}$ which provide solutions to Problem 1. Therefore, we choose the one which maximizes the probability of non-failure $P(T > x)$. The latter is achieved by finding an appropriate t_J, i.e. the moment of capital injection of size J. It is observed that, moving the location t_J of the jump $J = 2.5$ from $t_J = 0$ to $t_J = 2$ with the other parameter values fixed, the survival probability $P(T > 2)$ first increases slowly and then drops quickly, and a unique maximum is reached at $t_J = 0.9$.

It has to be mentioned that the numerical illustration in this example is rather simple, based on the specific class of functions $\mathcal{H}$ selected, and a unique solution to Problem 1 of maximizing the finite-time survival probability has been obtained. With more general choices of $\mathcal{H}$ in more complex situations

in system reliability risk analysis, one may not be able to find an optimal solution to Problem 1 and a more involved functional analysis approach may need to be utilized. The framework itself may also need to be calibrated carefully in order to fit the real scenarios properly, but the methodology described in section 2 would be general enough to cover most, if not all, real system characteristics.

3.2.2 Inventory management risk

This example illustrates the application of our risk modelling approach to management and maintenance of an inventory of products over a predetermined time period. There are two important processes related to an inventory of a product, the demand process which is described by the amount of units and time instants they are withdrawn from inventory and the replenishment process which specifies the amounts of units and instants at which they are replenished to inventory. It is essential to note that in real inventory problems considerable uncertainty may be related to both the times of arrivals and levels of orders for needed production materials and finished goods (see e.g. Hillier and Lieberman 2010, Ch. 18). This leads to the necessity to construct stochastic inventory maintenance models in which both demand and/or replenishment of inventory follow appropriate stochastic processes. We note that the risk analysis approaches described above under both models A and B, can be successfully applied in modelling the amount of products maintained in inventory so that the probability of demand exceeding supply, i.e., the probability of a shortage in inventory over a fixed time interval, is kept sufficiently low. Under model A it would be natural to assume that predetermined fixed amounts of units are stocked in inventory at certain predetermined dates, which in aggregate forms the system's (deterministic) gain process $h(t)$, as described in section 2.1. The demand process $S(t)$ is considered stochastic, assuming that certain orders for random quantities $W_1, W_2, \ldots$ of units arrive at random times $T_1, T_2, \ldots$ following e.g. market needs. This is in line with $S(t)$ given in (1) (see also the assumptions of model A in section 2.1). The probability of demand exceeding supply within a finite time interval of size x can then be evaluated by utilizing formula (3). In this case, it is relevant to consider as failure the event $\{T < x, Y > y\}$, where T is the time the demand process exceeds the replenishment process, and given this, the deficit in demand Y exceeds a pre-determined acceptable level y, and formula (5) can then be applied to compute the likelihood of this event. Alternatively, in many real inventory maintenance situations, it is relevant to assume that inversely, the demand process follows a deterministic pattern $h(t)$, e.g. orders for fixed predetermined volumes of goods are ordered to be shipped at predetermined dates, as in the assumptions of model B, while the

replenishment of goods follows a stochastic process, $S(t)$, as specified in section 2.2. There are various examples of such inventory management models in the literature, see e.g. Heyman and Sobel (2003); Resurreccion and Santos (2012).

3.2.3 Risk of emerging disease (ED) spread losses

In this application, losses occur due to an emerging (infectious) disease (ED) affecting e.g. fish populations in catfish farms in US. A risk analytic approach to modelling such losses at a single farm level, was recently proposed by Zagmutt et al. (2013), for the purpose of assessing the possibility of launching agricultural insurance cover to protect catfish producers. The authors apply stochastic disease spread modelling (DSM), implemented via Monte Carlo (MC) simulation in order to estimate the total loss cost, LC_{ed} (see Eq. (1) in Zagmutt et al. (2013)), due to fish mortality resulting from the spread of an ED in fish ponds belonging to a farm.

This approach can be further extended to model the losses from disease spread at a multi-farm level, where we focus on infection spreading between farms rather than between ponds within a single farm. Following Zagmutt et al. (2013), it is again assumed that disease spread between farms can be triggered by direct and indirect contacts, modelled by a Poisson process. The numerical results in Zagmutt et al. (2013) reveal that the total loss caused by fish mortalities for a single farm, LC_{ed}, is a random variable which follows a certain distribution, that is estimated using MC simulation. Thus, in a multi-farm system, where an ED spreads from one farm to another, the loss from each individual farm, LC_{ed}, can be modelled by a random variable W_i, $i = 1, 2, \ldots$. Therefore, it is natural to assume a certain dependence structure governing the farm loss severities, $W_1, W_2, \ldots$, as specified in our model A of section 2.1. Some organizations (or certain lines of business in an insurance company) may be concerned with the total loss caused by disease spread between farms at a state or national level, and may therefore be interested in providing some protection (e.g. insurance) for the total loss against an appropriate cumulative premium $h(t)$. Such a problem setup naturally meets the assumptions of model A described in section 2.1, and the results on the probability that the total loss from fish mortality at a multi-farm level exceeds the accumulated premium, $h(t)$, can be therefore applied in this case in order to determine whether this risk is insurable.

3.2.4 Risk of financial ruin and insolvency

Our final application is concerned with the risk of financial ruin and insolvency. The event of a technical ruin (insolvency), i.e. failure, occurs when the aggregate expense outgo exceeds the aggregate income inflow within a predetermined finite time horizon. Modelling the risk of ruin and insolvency and allocating sufficient risk capital, so as to minimize the probability of insolvency, is crucial for the healthy financial operation of firms. Companies from the financial services sector, such as insurance companies, hedge funds and other investment funds and banks are exposed to the occurrence of liabilities, typically insurance claims, devaluation of assets, credit defaults etc., which occur at some random moments in time τ_i and whose sizes are also considered random variables W_i as in the assumptions of model A (see section 2.1). On the other hand there is an inflow of income from operations of such companies. Typically, this would be premium income, capital gains from interests on deposits, mortgages, investments or other financial transactions, as specified in the assumptions of Model A. The probability of technical ruin within a finite-time horizon can be evaluated and used to determine the size of the risk capital. For further details of how model A is applied in the context of operational risk of financial firms we refer to Kaishev et al. (2008).

Alternatively, one can apply model B for other non-financial companies, where the W's are interpreted as capital gains occurring as a result of successful discoveries or mineral finds, or innovations, while aggregate operational costs follow a non-decreasing deterministic function (see section 2.2). For specific details of how model B is employed in modelling the operations of companies with continuous expenses and stochastic capital gains we refer to Dimitrova et al. (2013b).

4 DISCUSSION AND CONCLUSIONS

We have introduced two dually connected stochastic models A and B for quantifying the risk of failure and have interpreted the related model parameters and variables to fit specific applications in risk analysis. Under very general assumptions, we have provided new expressions for the probability that a risk factor exceeds a possibly time-dependent risk level, while at the same time, the exceedance over this level is greater than a pre-determined (deficit) value. Thus, we have demonstrated that expressions (3), (5) and (12) are quite general and hence, applicable in providing solutions to many risk analytic problems appearing in dam management, systems reliability, inventory management, infectious disease spread and financial insolvency (see section 3). We have further demonstrated

that having such explicit expressions is numerically appealing in quantifying the risk of failure and designing early warning systems incorporating alarm times. We note that the definition of an alarm time is not unique. As suggested by Das and Kratz (2010), who devise various alarm systems in the context of actuarial science and insurance, alternative ways of defining alarm times may be pursued, see also Dimitrova et al. (2014). Finally, let us also note that in some real situations it may also be necessary to compute the infinite-time horizon failure probability. For instance, failure and reliability of a system may be of interest over a very long (infinite) time-horizon. Such probabilities can be determined using e.g. the results due to Ignatov and Kaishev (2006).

References

Asmussen, S., H. Albrecher. 2010. *Ruin Probabilities*, World Scientific, New Jersey, 2nd ed.

Basili, M., Fanzini, M. 2006. Understanding the Risk of an Avian Flu Pandemic: Rational Waiting or Precautionary Failure? *Risk Analysis, 26*(3), 617–630.

Bullard, K.L. 1986. *Comparison of estimated probable maximum flood peaks with historic floods*. Hydrology Branch, Division of Planning Technical Services, Engineering and Research Center, U.S. Department of the Interior, Bureau of Reclamation.

Cheng, S., Yen, B.C., Tang, W.H. 1993. Stochastic Risk Modeling of Dam Overtopping. *Reliability and Uncertainty Analyses in Hydraulic Design: A Report*. American Society of Civil Engineers. 123–132.

Cox, L.A., Jr. 2009. *Risk Analysis of Complex and Uncertain Systems*. Springer, New York.

Das, S., Kratz, M. 2010. *On devising various alarm systems for insurance companies*. IIM Bangalore Working Paper 332 and ESSEC Working Paper 10008.

Das, S., Kratz, M. 2012. Alarm system for insurance companies: A strategy for capital allocation, *Insurance: Mathematics and Economics, 51*(1), 53–65.

Dimitrova, D.S., Kaishev, V.K., Zhao, S. 2013a. On the evaluation of finite-time ruin probabilities in a dependent risk model. *Submitted.* (Available at SSRN: http://ssrn.com/abstract=2432454)

Dimitrova, D.S., Kaishev, V.K., Zhao, S. 2013b. On finite-time ruin probabilities in a generalized dual risk model with dependence. *European Journal of Operational Research,* forthcoming.

Dimitrova, D.S., Kaishev, V.K., Zhao, S. 2014. Early warning of bankruptcy and risk capital allocation based on the time to ruin and the deficit at ruin. *Submitted.*

Emily, M., Casez, P. and Francois, O. 2009. Risk Assessment for Hospital-Acquired Diseases: A Risk-Theory Approach. *Risk Analysis, 29*(4), 565–575.

Gerrard, R., Tsanakas, A. 2011. Failure Probability Under Parameter Uncertainty. *Risk Analysis, 31*(5), 727–744.

Goodarzi, E., Mirzaei, M., Shui, L.T., Ziaei, M. 2011. Evaluation dam overtopping risk based on univariate and bivariate flood frequency analysis. *Hydrology and Earth System Sciences Discussions, 8*, 9757–9796.

Heyman, D.P., Sobel, M.J. 2003. *Stochastic Models in Operations Research: Stochastic Processes and Operating Characteristics*. Vol. I. Dover Publications.

Hillier, F.S., Lieberman, G.J. 2010. *Introduction to operations research*. 9th edition. McGraw-Hill, New York.

Ignatov, Z.G., Kaishev, V.K. 2000. Two-sided Bounds for the Finite-time Probability of Ruin. *Scandinavian Actuarial Journal, 2000*(1), 46–62.

Ignatov, Z.G., Kaishev, V.K. 2004. A Finite-Time Ruin Probability Formula For Continuous Claim Severites. *Journal of Applied Probability, 41*(2), 570–578.

Ignatov, Z.G., Kaishev, V.K. 2006. On the infinite-horizon probability of (non)ruin for integer-valued claims. *Journal of Applied Probability, 43*(2), 535–551.

Ignatov, Z.G., Kaishev, V.K. 2012. *On the probability of ruin and the deficit at ruin*. Paper presented at Gnedenko 100 Conference: International Conference on Probability Theory and its Applications, Moscow, June 2012.

Ignatov, Z.G., Kaishev, V.K. 2012. Finite time non-ruin probability for Erlang claim inter-arrivals and continuous inter-dependent claim amounts. *Stochastics: An International Journal of Probability and Stochastic Processes, 84*(4), 461–485.

Ignatov, Z.G., Kaishev, V.K., Krachunov, R.S. 2001. An Improved Finite-time Ruin Probability Formula and its "Mathematica" Implementation. *Insurance: Mathematics and Economics, 29*(3), 375–386.

Kaishev, V.K., Dimitrova, D.S., Ignatov, Z.G. 2008. Operational Risk and Insurance: A Ruin-probabilistic Reserving Approach. *Journal of Operational Risk, 3*(3), 39–60.

Kolen, B., Kok, M., Helsloot, I., Maaskant, B. 2013. EvacuAid: A Probabilistic Model to Determine the Expected Loss of Life for Different Mass Evacuation Strategies During Flood Threats. *Risk Analysis, 33*(7), 1312–1333.

Kwon, H., Moon, Y. 2006. Improvement of overtopping risk evaluations using probabilistic concepts for existing dams. *Stochastic Environmental Research and Risk Assessment, 20*, 223–237.

Landriault, D., Willmot, G.E. 2009. On the joint distributions of the time to ruin, the surplus prior to ruin, and the deficit at ruin in the classical risk model. *North American Actuarial Journal, 13*, 252–270.

Lave, L.B., Balvanyos, T. 1998. Risk analysis and management of dam safety. *Risk Analysis, 18*(4), 455–462.

Paté-Cornell, M.E. 1986. Warning Systems in Risk Management. *Risk Analysis, 6*(2), 223–234.

Paté-Cornell, M.E., Bea, R.G. 1992. Management Errors and System Reliability: A Probabilistic Approach and Application to Offshore Platforms. *Risk Analysis, 12*(1), 1–18.

Resurreccion, J. and Santos, J.R. 2012. Multiobjective Prioritization Methodology and Decision Support System for Evaluating Inventory Enhancement Strategies for Disrupted Interdependent Sectors. *Risk Analysis, 32*(10), 1673–1692.

Serrano-Lombillo, A., Escuder-Bueno, I., de Membrillera-Ortuno, M.G., Altarejos-García, L. 2011. Methodology for the Calculation of Annualized Incremental Risks in Sytems of Dams. *Risk Analysis, 31*(6), 1000–1015.

Tingsanchali, T., Chinnarasri, C. 2001. Numerical modelling of dam failure due to flow overtopping. *Hydrological Sciences Journal, 46*(1), 113–130.

Todinov, M.T. 2006. Reliability analysis based on the losses from failures. *Risk Analysis, 26*(2), 311–335.

Valdes, J.B., Marco, J.B. 1995. *Manageing reservoirs for flood control.* U.S.–Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods. Perugia (Italy), November 1995.

Zagmutt, F.J., Sempier, S.H., Hanson, T.R. 2013. Disease spread models to estimate highly uncertain emerging diseases losses for animal agriculture insurance policies: an application to the U.S. farm-raised catfish industry. *Risk Analysis, 33*(10), 1924–1936.

Appendix

A A proof on the equivalence of formula (3) with (5) when $y = 0$

In this section, we provide a detailed proof which shows that the expression of $P(T > x)$ given in formula (3) coincides with the expression of $P(T < x, Y > y)$ given in (5) in the case where $y = 0$.

Proof: When $y = 0$, rearranging the terms in formula (5), we have

$$\begin{aligned}
P(T < x, Y > 0) &\equiv P(T < x) \\
&= 1 + \sum_{k=2}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{y_{k-1}}^{\infty} B_{k-2}\Big(h^{-1}(y_{k-1}); y_1, \ldots, y_{k-2}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&\quad - \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{y_{k-1}}^{h(x)} B_{k-1}\Big(h^{-1}(y_k); y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&\quad - \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{h(x)}^{\infty} B_{k-1}\Big(x; y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
P(T > x) &= -\sum_{k=2}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{y_{k-1}}^{\infty} B_{k-2}\Big(h^{-1}(y_{k-1}); y_1, \ldots, y_{k-2}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&\quad + \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{y_{k-1}}^{h(x)} B_{k-1}\Big(h^{-1}(y_k); y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&\quad + \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{h(x)}^{\infty} B_{k-1}\Big(x; y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= -① + ② + ③.
\end{aligned}$$

We then have

$$\begin{aligned}
① &= \sum_{k=2}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{y_{k-1}}^{\infty} B_{k-2}\Big(h^{-1}(y_{k-1}); y_1, \ldots, y_{k-2}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= \sum_{k=2}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} B_{k-2}\Big(h^{-1}(y_{k-1}); y_1, \ldots, y_{k-2}\Big) \Big(\int_{y_{k-1}}^{\infty} f(y_1, \ldots, y_k) dy_k \Big) dy_{k-1} \ldots dy_1 \\
&= \sum_{k=2}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} B_{k-2}\Big(h^{-1}(y_{k-1}); y_1, \ldots, y_{k-2}\Big) f(y_1, \ldots, y_{k-1}) dy_{k-1} \ldots dy_1 \\
&= \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-1}}^{h(x)} B_{k-1}\Big(h^{-1}(y_k); y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= ②,
\end{aligned}$$

indicating that

$$P(T > x) \quad = \quad -① + ② + ③ = ③.$$

Thus,

$$\begin{aligned}
P(T > x) &= \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{h(x)}^{\infty} B_{k-1}\Big(x; y_1, \ldots, y_{k-1}\Big) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{h(x)}^{\infty} e^{-\lambda x} \sum_{i=0}^{k-1} A_i(\lambda x; \lambda\nu_1, \ldots, \lambda\nu_i) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= e^{-\lambda x} \sum_{k=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-2}}^{h(x)} \int_{h(x)}^{\infty} \sum_{i=0}^{k-1} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) f(y_1, \ldots, y_k) dy_k \ldots dy_1 \\
&= e^{-\lambda x} \sum_{k=0}^{\infty} \sum_{i=0}^{k} \int_0^{h(x)} \cdots \int_{y_{k-1}}^{h(x)} \int_{h(x)}^{\infty} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) f(y_1, \ldots, y_{k+1}) dy_{k+1} \ldots dy_1.
\end{aligned}$$

Permuting the two sums, we obtain

$$\begin{aligned}
P(T > x) &= e^{-\lambda x} \sum_{i=0}^{\infty} \sum_{k=i}^{\infty} \int_0^{h(x)} \cdots \int_{y_{k-1}}^{h(x)} \int_{h(x)}^{\infty} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) f(y_1, \ldots, y_{k+1}) dy_{k+1} \ldots dy_1 \\
&= e^{-\lambda x} \sum_{i=0}^{\infty} \int_0^{h(x)} \cdots \int_{y_{i-1}}^{h(x)} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) \times \\
&\quad \left(\sum_{k=i}^{\infty} \int_{y_i}^{h(x)} \cdots \int_{y_{k-1}}^{h(x)} \int_{h(x)}^{\infty} f(y_1, \ldots, y_{k+1}) dy_{k+1} \ldots dy_{i+1} \right) dy_i \ldots dy_1 \\
&= e^{-\lambda x} \sum_{i=0}^{\infty} \int_0^{h(x)} \cdots \int_{y_{i-1}}^{h(x)} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) \times \\
&\quad \Big(\int_{h(x)}^{\infty} f(y_1, \ldots, y_{i+1}) dy_{i+1} + \int_{y_i}^{h(x)} \int_{h(x)}^{\infty} f(y_1, \ldots, y_{i+2}) dy_{i+2} dy_{i+1} \\
&\quad \int_{y_i}^{h(x)} \int_{y_{i+1}}^{h(x)} \int_{h(x)}^{\infty} f(y_1, \ldots, y_{i+3}) dy_{i+3} dy_{i+2} dy_{i+1} + \cdots \Big) dy_i \ldots dy_1.
\end{aligned}$$

Noting that the term in the brackets is identically equal to $f(y_1, \ldots, y_i)$, we have

$$\begin{aligned}
P(T > x) &= e^{-\lambda x} \sum_{i=0}^{\infty} \int_0^{h(x)} \cdots \int_{y_{i-1}}^{h(x)} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) f(y_1, \ldots, y_i) dy_i \ldots dy_1 \\
&= e^{-\lambda x} \left(1 + \sum_{i=1}^{\infty} \int_0^{h(x)} \cdots \int_{y_{i-1}}^{h(x)} \lambda^i A_i(x; \nu_1, \ldots, \nu_i) f(y_1, \ldots, y_i) dy_i \ldots dy_1 \right),
\end{aligned}$$

which completes the proof.

B A derivation of the joint probability of the time to ruin and the deficit at ruin in model A assuming independent, non-identically Erlang distributed inter-arrival times

Here, we provide a detailed derivation of the joint probability that the time to ruin $T < x$ and the deficit at ruin $Y > y$ in model A with independent non-identically Erlang distributed inter-arrival times. In what follows, we only consider the case of continuous losses. However, we note that, for the case of discrete losses, an analogous derivation can be followed and similar results could be obtained.

Under the setup of model A, let τ_i, $i = 1, 2, \ldots$, be a sequence of independent, non-identical random variables, denoting the inter-arrival times. We assume these are independent, non-identically Erlang distributed random variables with shape parameter $g_i > 0$, and rate parameter $\lambda_i > 0$, i.e. $\tau_i \sim$Erlang(g_i, λ_i). Let $l_i = g_1 + \cdots + g_i$, $i = 1, 2, \ldots$. Similar to Ignatov and Kaishev (2012b), we introduce an integer-valued function $j(k)$, $k = 0, 1, 2, \ldots$, such that

$$g_1 + \ldots + g_{j(k)} = l_{j(k)} \leq k < l_{j(k)+1} = g_1 + \ldots + g_{j(k)} + g_{j(k)+1}. \tag{10}$$

Let $\{\tilde{\tau}_n\}_{n\geq 1}$ be a sequence of independent, exponentially distributed random variables with parameters $\theta_1, \theta_2, \ldots$ correspondingly, i.e. $\tilde{\tau}_n \sim$Exp(θ_n), such that $\theta_{k+1} = \lambda_{j(k)+1}$, $k = 0, 1, 2, \ldots$. Thus, we have

$$(\tilde{\tau}_1 + \cdots + \tilde{\tau}_{l_1}, \tilde{\tau}_{l_1+1} + \cdots + \tilde{\tau}_{l_2}, \ldots) \stackrel{d}{=} (\tau_1, \tau_2, \ldots).$$

Obviously, in this more refined representation of the Erlang arrivals in terms of sums of exponentials we have that

$$\theta_1, \ldots, \theta_{l_1}, \theta_{l_1+1}, \ldots, \theta_{l_2}, \ldots \equiv \underbrace{\lambda_1, \ldots, \lambda_1}_{g_1}, \underbrace{\lambda_2, \ldots, \lambda_2}_{g_2}, \ldots, \tag{11}$$

noting that the λ_i-s may possibly coincide. In the sequel, it will be convenient to use the notation $\tilde{\tau}_1^*, \tilde{\tau}_2^*, \ldots$ for the r.v.s $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$, in the special case when $\theta_{k+1} = \lambda_{j(k)+1} \equiv 1$, $k = 0, 1, 2, \ldots$.

Denote by $T_1 = \tau_1, T_2 = \tau_1 + \tau_2, \ldots$, the moments of loss arrivals and introduce the sequence of random variables $\tilde{T}_1 = \tilde{\tau}_1$, $\tilde{T}_2 = \tilde{\tau}_1 + \tilde{\tau}_2, \ldots$. Obviously, we can also write $T_i = \tilde{T}_{l_i}$ $i = 1, 2, \ldots$. Let us also consider the partial sums, Y_i, $i = 1, 2, \ldots$ of the consecutive losses, $Y_1 = W_1, Y_2 = W_1 + W_2, \ldots$

with probability density function $f_{Y_1,\ldots,Y_i}(y_1,\ldots,y_i) \geq 0$ for $0 \leq y_1 \leq \ldots \leq y_i$ and 0 otherwise, i.e.

$$\int\limits_{0\leq y_1\leq\ldots\leq y_i} \ldots \int f_{Y_1,\ldots,Y_i}(y_1,\ldots,y_i)dy_1\ldots dy_i = 1.$$

We will also denote by $F_{Y_1,\ldots,Y_i}(y_1,\ldots,y_i)$, the cdf of $Y_1,\ldots,Y_i$.

We now introduce the non-decreasing sequence of variables $\tilde{Y}_1, \tilde{Y}_2, \ldots$, independent of $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$ and such that $0 = \tilde{Y}_1 = \ldots = \tilde{Y}_{l_1-1} \leq Y_1 = \tilde{Y}_{l_1} = \ldots = \tilde{Y}_{l_2-1} \leq Y_2 = \tilde{Y}_{l_2} = \ldots = \tilde{Y}_{l_3-1} \leq \ldots$.

Then we have the following proposition.

Proposition B.1 *The probability $P(T < x, Y > y)$, $x > 0$, $y \geq 0$, is given by (assuming $l_1 \geq 2$)*

$$\begin{aligned}
&P(T < x, Y > y) = \\
&\sum_{k=1}^{\infty} \int\limits_0^{h(x)} \ldots \int\limits_{y_{k-2}}^{h(x)} \int\limits_{y_{k-1}+y}^{\infty} \mathfrak{B}_{l_k-2}\left(h^{-1}(y_{k-1});\nu_1,\ldots,\nu_{l_k-2}\right) f_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k)dy_k\ dy_{k-1}\ldots dy_1 \\
&-\sum_{k=1}^{\infty} \int\limits_0^{h(x)} \ldots \int\limits_{y_{k-2}}^{h(x)} \int\limits_{y_{k-1}+y}^{h(x)+y} \mathfrak{B}_{l_k-1}\left(h^{-1}(y_k - y);\nu_1,\ldots,\nu_{l_k-1}\right) f_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k)dy_k\ dy_{k-1}\ldots dy_1 \\
&-\sum_{k=1}^{\infty} \int\limits_0^{h(x)} \ldots \int\limits_{y_{k-2}}^{h(x)} \int\limits_{h(x)+y}^{\infty} \mathfrak{B}_{l_k-1}\left(x;\nu_1,\ldots,\nu_{l_k-1}\right) f_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k)dy_k\ dy_{k-1}\ldots dy_1, \qquad (12)
\end{aligned}$$

where

$$\mathfrak{B}_k(z;\nu_1,\ldots,\nu_k) = \sum_{i=0}^{k} B_k(z;\nu_1,\ldots,\nu_k),$$

$B_k(z;\nu_1,\ldots,\nu_k) \equiv B_k(z)$ are the (classical) exponential Appell polynomials, defined by Ignatov and Kaishev (2012b) recurrently as

$$B_k(z) = \lambda_{j(k-1)+1}e^{-\lambda_{j(k)+1}z}\int_{\nu_{j(k)}}^{z} e^{\lambda_{j(k)+1}w}B_{k-1}(w)dw, k = 1,2,\ldots$$

with $B_0(z) = \mathfrak{B}_0(z) = e^{-\lambda_1 z}$ and $0 \leq \nu_1 \leq \nu_2 \leq \ldots$ is a sequence of real numbers denoting

$$\underbrace{h^{-1}(0) \leq \ldots \leq h^{-1}(0)}_{g_1-1} \leq \underbrace{h^{-1}(y_1) \leq \ldots \leq h^{-1}(y_1)}_{g_2} \leq \ldots,$$

correspondingly.

Proof: By construction, the event $\{T < x, Y > y\}$ can be expressed as

$$\begin{aligned}\{T < x, Y > y\} &= \bigcup_{k=1}^{\infty}\left[\bigcap_{i=1}^{k-1}\{Y_i < h(T_i)\} \cap \{Y_k > h(T_k) + y\} \cap \{T_k < x\}\right] \\ &= \bigcup_{k=1}^{\infty}\left[\bigcap_{i=1}^{k-1}\left\{\tilde{Y}_{l_i} < h\left(\tilde{T}_{l_i}\right)\right\} \cap \left\{\tilde{Y}_{l_k} > h\left(\tilde{T}_{l_k}\right) + y\right\} \cap \left\{\tilde{T}_{l_k} < x\right\}\right].\end{aligned}$$

Clearly, $\left\{\tilde{Y}_{l_i} < h\left(\tilde{T}_{l_i}\right)\right\} \subseteq \left\{\tilde{Y}_{l_i} < h\left(\tilde{T}_{l_i+w}\right)\right\}$ for $w = 0, 1, \ldots, g_{i+1} - 1$, which is equivalent to $\left\{\tilde{Y}_{l_i} < h\left(\tilde{T}_{l_i}\right)\right\} \subseteq \left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\}$ for $l_i \leq r < l_{i+1}$. Therefore, for any $i = 1, 2, \ldots,$

$$\left\{\tilde{Y}_{l_i} < h\left(\tilde{T}_{l_i}\right)\right\} \subseteq \bigcap_{r=l_i}^{l_{i+1}-1}\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\}.$$

In addition, for $1 \leq r < l_1$, we also have $\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\} = \left\{0 < h\left(\tilde{T}_r\right)\right\} = \Omega$, and hence $\bigcap_{r=1}^{l_1-1}\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\} = \Omega$, where Ω is the sure event. Thus, we obtain

$$\{T < x, Y > y\} = \bigcup_{k=1}^{\infty}\left[\bigcap_{r=1}^{l_k-1}\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\} \cap \left\{\tilde{Y}_{l_k} > h\left(\tilde{T}_{l_k}\right) + y\right\} \cap \left\{\tilde{T}_{l_k} < x\right\}\right].$$

We note that the events in the square brackets are mutually exclusive. Hence, we have

$$\begin{aligned}& P(T < x, Y > y) \\ = \;& P\left(\bigcup_{k=1}^{\infty}\left[\bigcap_{r=1}^{l_k-1}\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\} \cap \left\{\tilde{Y}_{l_k} > h\left(\tilde{T}_{l_k}\right) + y\right\} \cap \left\{\tilde{T}_{l_k} < x\right\}\right]\right) \\ = \;& \sum_{k=1}^{\infty} P\left(\bigcap_{r=1}^{l_k-1}\left\{\tilde{Y}_r < h\left(\tilde{T}_r\right)\right\} \cap \left\{\tilde{Y}_{l_k} > h\left(\tilde{T}_{l_k}\right) + y\right\} \cap \left\{\tilde{T}_{l_k} < x\right\}\right) \qquad (13) \\ = \;& \sum_{k=1}^{\infty}\int_0^x \cdots \int_{t_{l_k-1}}^x f_{\tilde{T}_1,\ldots,\tilde{T}_{l_k}}(t_1, \ldots, t_{l_k}) dt_{l_k} \ldots dt_1 \int_0^{h(t_1)} \cdots \int_{\tilde{y}_{l_k-2}}^{h(t_{l_k-1})} \int_{h(t_{l_k})+y}^{\infty} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}),\end{aligned}$$

where $F_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k})$ is the joint distribution of $\tilde{Y}_1, \ldots, \tilde{Y}_{l_k}$ and $f_{\tilde{T}_1,\ldots,\tilde{T}_{l_k}}(t_1, \ldots, t_{l_k})$ is the joint density of $\tilde{T}_1, \ldots, \tilde{T}_{l_k}$. It can easily be seen that the random vector $\tilde{\boldsymbol{T}} = \left(\tilde{T}_1, \ldots, \tilde{T}_{l_k}\right)'$ coincides in distribution with the random vector $\boldsymbol{B}_{l_k}\tilde{\boldsymbol{\tau}}^*$, i.e., $\boldsymbol{B}_{l_k}\tilde{\boldsymbol{\tau}}^* \stackrel{d}{=} \tilde{\boldsymbol{T}}$, where $\tilde{\boldsymbol{\tau}}^* = \left(\tilde{\tau}_1^*, \ldots, \tilde{\tau}_{l_k}^*\right)'$, and $\boldsymbol{B}_{l_k}$ is a

$l_k \times l_k$ dimensional matrix given in a block-matrix form as

$$\boldsymbol{B}_{l_k} \equiv \begin{pmatrix} \boldsymbol{b}_{1,1} & \cdots & \boldsymbol{b}_{1,k} \\ \vdots & \ddots & \vdots \\ \boldsymbol{b}_{k,1} & \cdots & \boldsymbol{b}_{k,k} \end{pmatrix},$$

where $\boldsymbol{b}_{m,n}$ is a $g_m \times g_n$ matrix for $m, n = 1, \ldots, k$, with all entries equal to $\frac{1}{\lambda_n}$ if $m > n$, all entries equal to zero if $m < n$, and where $\boldsymbol{b}_{n,n}$ is a lower triangular matrix with all elements in the lower triangle equal to $\frac{1}{\lambda_n}$ if $m = n$. Then, it is not difficult to see that

$$f_{\tilde{T}_1,\ldots,\tilde{T}_{l_k}}(t_1, \ldots, t_{l_k}) = \begin{cases} e^{-\mathbf{1}\cdot\boldsymbol{B}_{l_k}^{-1}\cdot\boldsymbol{t}} \left|\det \boldsymbol{B}_{l_k}^{-1}\right| & \text{if } 0 \le t_1 \le t_2 \le \ldots \le t_{l_k} \\ 0 & \text{otherwise} \end{cases},$$

where, $\mathbf{1} = \left(\underbrace{1, \ldots, 1}_{l_k}\right)$, $\boldsymbol{t} = (t_1, \ldots, t_{l_k})'$, $()'$ stands for transposition, and $\det \boldsymbol{B}_{l_k}^{-1}$ denotes the determinant of the inverse of $\boldsymbol{B}_{l_k}$. It can also be directly verified that the inverse matrix, $\boldsymbol{B}_{l_k}^{-1} = \left(\tilde{b}_{ij}\right)$, is an incomplete, lower triangular matrix, with non-zero elements only at the main and next lower diagonals, given as

$$\tilde{b}_{m,m} = \begin{cases} \lambda_1 & \text{if} \quad 1 \le m \le l_1 \\ \lambda_2 & \text{if} \quad l_1 + 1 \le m \le l_2 \\ \vdots & \vdots \quad\quad \vdots \\ \lambda_k & \text{if} \quad l_{k-1} + 1 \le m \le l_k \end{cases}, \tilde{b}_{m+1,m} = \begin{cases} -\lambda_1 & \text{if} \quad 1 \le m \le l_1 - 1 \\ -\lambda_2 & \text{if} \quad l_1 \le m \le l_2 - 1 \\ \vdots & \vdots \quad\quad \vdots \\ -\lambda_k & \text{if} \quad l_{k-1} \le l \le l_k - 1 \end{cases},$$

and with all other elements equal to zero. Then, $P(T < x, Y > y)$ becomes

$$\begin{aligned} & P(T < x, Y > y) \\ = & \sum_{k=1}^{\infty} \int_0^x \cdots \int_{t_{l_k-1}}^x \lambda_1^{g_1} \ldots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2} - t_{l_1}) + \cdots + \lambda_k(t_{l_k} - t_{l_k-1})\}] dt_{l_k} \ldots dt_1 \\ & \int_0^{h(t_1)} \cdots \int_{\tilde{y}_{l_k-2}}^{h(t_{l_k-1})} \int_{h(t_{l_k})+y}^{\infty} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}). \end{aligned}$$

Splitting the innermost integral, we have

$$
\begin{aligned}
& P(T<x, Y>y) \\
= & \sum_{k=1}^{\infty} \int_0^x \cdots \int_{t_{l_{k-1}}}^x \lambda_1^{g_1} \ldots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2}-t_{l_1}) + \cdots + \lambda_k(t_{l_k}-t_{l_{k-1}})\}] dt_{l_k} \ldots dt_1 \\
& \int_0^{h(t_1)} \cdots \int_{\tilde{y}_{l_k-2}}^{h(t_{l_k-1})} \left(\int_{\tilde{y}_{l_k-1}+y}^{\infty} - \int_{\tilde{y}_{l_k-1}+y}^{h(t_{l_k})+y} \right) dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}) \\
= & \sum_{k=1}^{\infty} \int_0^x \cdots \int_{t_{l_k-2}}^x \lambda_1^{g_1} \ldots \lambda_{k-1}^{g_{k-1}} \lambda_k^{g_k-1} \Big(\exp[-\{\lambda_1 t_{l_1} + \cdots + \lambda_{k-1}(t_{l_{k-1}} - t_{l_{k-2}}) + \lambda_k(t_{l_k-1} - t_{l_{k-1}})\}] \\
& - \exp[-\{\lambda_1 t_1 + \cdots + \lambda_{k-1}(t_{l_{k-1}} - t_{l_{k-2}}) + \lambda_k(x - t_{l_{k-1}})\}] \Big) dt_{l_k-1} \ldots dt_1 \\
& \int_0^{h(t_1)} \cdots \int_{\tilde{y}_{l_k-2}}^{h(t_{l_k-1})} \int_{\tilde{y}_{l_k-1}+y}^{\infty} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}) \\
& - \sum_{k=1}^{\infty} \int_0^x \cdots \int_{t_{l_k-1}}^x \lambda_1^{g_1} \ldots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2}-t_{l_1}) + \cdots + \lambda_k(t_{l_k}-t_{l_{k-1}})\}] dt_{l_k} \ldots dt_1 \\
& \int_0^{h(t_1)} \cdots \int_{\tilde{y}_{l_k-2}}^{h(t_{l_k-1})h(t_{l_k})+y} \int_{\tilde{y}_{l_k-1}+y} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}).
\end{aligned}
$$

Permuting the two multiple integrals in the two terms, we obtain

$$
\begin{aligned}
& P(T<x, Y>y) \\
= & \sum_{k=1}^{\infty} \int \cdots \int_{\substack{0<\tilde{y}_1<\ldots<\tilde{y}_{l_k-1}<h(x) \\ \tilde{y}_{l_k-1}+y<\tilde{y}_{l_k}<\infty}} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}) \int_{h^{-1}(\tilde{y}_1)}^{t_2} \cdots \int_{h^{-1}(\tilde{y}_{l_k-2})}^{t_{l_k-1}} \int_{h^{-1}(\tilde{y}_{l_k-1})}^{x} \lambda_1^{g_1} \ldots \lambda_{k-1}^{g_{k-1}} \lambda_k^{g_k-1} \\
& \Big(\exp[-\{\lambda_1 t_{l_1} + \cdots + \lambda_{k-1}(t_{l_{k-1}} - t_{l_{k-2}}) + \lambda_k(t_{l_k-1} - t_{l_{k-1}})\}] - \\
& \exp[-\{\lambda_1 t_1 + \cdots + \lambda_{k-1}(t_{l_{k-1}} - t_{l_{k-2}}) + \lambda_k(x - t_{l_{k-1}})\}] \Big) dt_{l_k-1} \ldots dt_1 \\
& - \sum_{k=1}^{\infty} \sum_{k=1}^{\infty} \int \cdots \int_{\substack{0<\tilde{y}_1<\ldots<\tilde{y}_{l_k-1}<h(x) \\ \tilde{y}_{l_k-1}+y<\tilde{y}_{l_k}<h(x)+y}} dF_{\tilde{Y}_1,\ldots,\tilde{Y}_{l_k}}(\tilde{y}_1, \ldots, \tilde{y}_{l_k}) \int_{h^{-1}(\tilde{y}_1)}^{t_2} \cdots \int_{h^{-1}(\tilde{y}_{l_k-1})}^{t_{l_k}} \int_{h^{-1}(\tilde{y}_{l_k}-y)}^{x} \lambda_1^{g_1} \ldots \lambda_k^{g_k} \\
& \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2}-t_{l_1}) + \cdots + \lambda_k(t_{l_k}-t_{l_{k-1}})\}] dt_{l_k} \ldots dt_1,
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
& P(T<x, Y>y) \\
= & \sum_{k=1}^{\infty} \int\cdots\int_{\substack{0<y_1<\ldots<y_{k-1}<h(x) \\ y_{k-1}+y<y_k<\infty}} dF_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k) \int_{\nu_1}^{t_2}\cdots\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{t_{l_k-1}}\int_{\nu_{l_k-1}h^{-1}(y_{k-1})}^{x} \lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1} \\
& \Big(\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]- \\
& \exp[-\{\lambda_1 t_1+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(x-t_{l_{k-1}})\}]\Big)dt_{l_k-1}\ldots dt_1 \\
& -\sum_{k=1}^{\infty} \int\cdots\int_{\substack{0<y_1<\ldots<y_{k-1}<h(x) \\ y_{k-1}+y<y_k<h(x)+y}} dF_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k) \int_{\nu_1}^{t_2}\cdots\int_{\nu_{l_k-1}h^{-1}(y_k-y)}^{t_{l_k}}\int_{\nu_{l_k-1}h^{-1}(y_k-y)}^{x} \lambda_1^{g_1}\ldots\lambda_k^{g_k} \\
& \exp[-\{\lambda_1 t_{l_1}+\lambda_2(t_{l_2}-t_{l_1})+\cdots+\lambda_k(t_{l_k}-t_{l_{k-1}})\}]dt_{l_k}\ldots dt_1 \\
= & \sum_{k=1}^{\infty}\int_0^{h(x)}\cdots\int_{y_{k-2}}^{h(x)}\int_{y_{k-1}+y}^{\infty} f_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k)dy_k\ dy_{k-1}\ldots dy_1 \int_{\nu_1}^{t_2}\cdots\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{t_{l_k-1}}\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{x} \lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1} \\
& \Big(\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]- \\
& \exp[-\{\lambda_1 t_1+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(x-t_{l_{k-1}})\}]\Big)dt_{l_k-1}\ldots dt_1 \\
& -\sum_{k=1}^{\infty}\int_0^{h(x)}\cdots\int_{y_{k-2}}^{h(x)}\int_{y_{k-1}+y}^{h(x)+y} f_{Y_1,\ldots,Y_k}(y_1,\ldots,y_k)dy_k\ dy_{k-1}\ldots dy_1 \int_{\nu_1}^{t_2}\cdots\int_{\nu_{l_k-1}h^{-1}(y_k-y)}^{t_{l_k}}\int_{\nu_{l_k-1}h^{-1}(y_k-y)}^{x} \lambda_1^{g_1}\ldots\lambda_k^{g_k} \\
& \exp[-\{\lambda_1 t_{l_1}+\lambda_2(t_{l_2}-t_{l_1})+\cdots+\lambda_k(t_{l_k}-t_{l_{k-1}})\}]dt_{l_k}\ldots dt_1, \qquad (14)
\end{aligned}
$$

where $0 \le \nu_1 \le \nu_2 \le \ldots$ is a sequence of real numbers denoting

$$
\underbrace{h^{-1}(0) \le \ldots \le h^{-1}(0)}_{g_1-1} \le \underbrace{h^{-1}(y_1) \le \ldots \le h^{-1}(y_1)}_{g_2} \le \ldots,
$$

correspondingly.

Let $\mathfrak{B}_0(z) = e^{-\theta_1 z}$ and for $k = 1, 2, \ldots$,

$$
\begin{aligned}
& \mathfrak{B}_k(z;\nu_1,\ldots,\nu_k) \equiv \mathfrak{B}_k(z) \\
= & \int_{\nu_1}^{t_2}\cdots\int_{\nu_k}^{t_{k+1}}\int_z^{\infty}\theta_1\ldots\theta_{k+1}\exp[-\{\theta_1 t_1+\theta_2(t_2-t_1)+\cdots+\theta_{k+1}(t_{k+1}-t_k)\}]dt_{k+1}\ dt_k\ldots dt_1,
\end{aligned}
$$

where $\{\theta_s\}_{s\ge1}$ follow the refined representation (11).

Permuting the two innermost integrals, we have

$$
\begin{aligned}
\mathfrak{B}_k(z) &= \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \left(\int_{\nu_k}^{t_{k+1}} \int_{z}^{\infty} \theta_1 \dots \theta_k \ \theta_{k+1} \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) + \theta_{k+1}(t_{k+1} - t_k)\}] \right. \\
&\quad \left. dt_{k+1} \ dt_k \right) dt_{k-1} \dots dt_1 \\
&= \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \left(\int_{\nu_k}^{\infty} \int_{\max\{t_k, z\}}^{\infty} \theta_1 \dots \theta_k \ \theta_{k+1} \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) - \theta_{k+1} t_k\}] \right. \\
&\quad \left. \times e^{-\theta_{k+1} t_{k+1}} dt_{k+1} \ dt_k \right) dt_{k-1} \dots dt_1 \\
&= \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \left(\int_{\nu_k}^{\infty} \theta_1 \dots \theta_k \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) - \theta_{k+1} t_k\}] \right. \\
&\quad \left. \times e^{-\theta_{k+1} \max\{t_k, z\}} dt_k \right) dt_{k-1} \dots dt_1 \\
&= \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \int_{\nu_k}^{z} \theta_1 \dots \theta_k \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) - \theta_{k+1} t_k\}] e^{-\theta_{k+1} z} dt_k \dots dt_1 \\
&\quad + \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \int_{z}^{\infty} \theta_1 \dots \theta_k \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) - \theta_{k+1} t_k\}] e^{-\theta_{k+1} t_k} dt_k \dots dt_1 \\
&= B_k(z; \nu_1, \dots, \nu_k) + \mathfrak{B}_{k-1}(z),
\end{aligned}
$$

where

$$
\begin{aligned}
& B_k(z; \nu_1, \dots, \nu_k) \equiv B_k(z) \\
= \ & e^{-\theta_{k+1} z} \int_{\nu_1}^{t_2} \dots \int_{\nu_{k-1}}^{t_k} \int_{\nu_k}^{z} \theta_1 \dots \theta_k \exp[-\{\theta_1 t_1 + \theta_2(t_2 - t_1) + \dots + \theta_k(t_k - t_{k-1}) - \theta_{k+1} t_k\}] dt_k \dots dt_1.
\end{aligned} \tag{15}
$$

Following Lemma A.1, A.2 and Corollary A.3 in Ignatov and Kaishev (2012b) and noting the refined representation (11), we have

$$
B_k(z) = \lambda_{j(k-1)+1} e^{-\lambda_{j(k)+1} z} \int_{\nu_{j(k)}}^{z} e^{\lambda_{j(k)+1} w} B_{k-1}(w) dw, k = 1, 2, \dots
$$

with $B_0(z) = e^{-\lambda_1 z}$ and $j(k)$ is as defined in (10). It can be easily seen that

$$\mathfrak{B}_k(z;\nu_1,\ldots,\nu_k) = \sum_{i=0}^{k} B_i(z;\nu_1,\ldots,\nu_i)$$

with $\mathfrak{B}_0(z) = B_0(z) = e^{-\theta_1 z}$.

Thus, we have

$$\begin{aligned}
&\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{t_{l_k-1}}\int^{x}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\Big(\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]\\
&-\exp[-\{\lambda_1 t_1+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(x-t_{l_{k-1}})\}]\Big)dt_{l_k-1}\ldots dt_1\\
=\;&\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{t_{l_k-1}}\int^{x}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]dt_{l_k-1}\ldots dt_1\\
&-\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}\nu_{l_k-1}}^{t_{l_k-1}}\int^{x}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\exp[-\{\lambda_1 t_1+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(x-t_{l_{k-1}})\}]dt_{l_k-1}\ldots dt_1\\
=\;&\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}h^{-1}(y_{k-1})}^{t_{l_k-1}}\int^{\infty}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]dt_{l_k-1}\ldots dt_1\\
&-\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}}^{t_{l_k-1}}\int_{x}^{\infty}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\exp[-\{\lambda_1 t_{l_1}+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(t_{l_k-1}-t_{l_{k-1}})\}]dt_{l_k-1}\ldots dt_1\\
&-\int_{\nu_1}^{t_2}\ldots\int_{\nu_{l_k-2}\nu_{l_k-1}}^{t_{l_k-1}}\int^{x}\lambda_1^{g_1}\ldots\lambda_{k-1}^{g_{k-1}}\lambda_k^{g_k-1}\exp[-\{\lambda_1 t_1+\cdots+\lambda_{k-1}(t_{l_{k-1}}-t_{l_{k-2}})+\lambda_k(x-t_{l_{k-1}})\}]dt_{l_k-1}\ldots dt_1\\
=\;&\mathfrak{B}_{l_k-2}\left(h^{-1}(y_{k-1});\nu_1,\ldots,\nu_{l_k-2}\right)-\mathfrak{B}_{l_k-2}\left(x;\nu_1,\ldots,\nu_{l_k-2}\right)-B_{l_k-1}\left(x;\nu_1,\ldots,\nu_{l_k-1}\right)\\
=\;&\mathfrak{B}_{l_k-2}\left(h^{-1}(y_{k-1});\nu_1,\ldots,\nu_{l_k-2}\right)-\mathfrak{B}_{l_k-1}\left(x;\nu_1,\ldots,\nu_{l_k-1}\right), \qquad (16)
\end{aligned}$$

noting $h^{-1}(y_{k-1}) = \nu_{l_k-1}$, and

$$
\begin{aligned}
& \int_{\nu_1}^{t_2} \dots \int_{\nu_{l_k-1} h^{-1}(y_k-y)}^{t_{l_k}} \int^{x} \lambda_1^{g_1} \dots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2} - t_{l_1}) + \dots + \lambda_k(t_{l_k} - t_{l_{k-1}})\}] dt_{l_k} \dots dt_1 \\
= & \int_{\nu_1}^{t_2} \dots \int_{\nu_{l_k-1} h^{-1}(y_{l_k}-y)}^{t_{l_k}} \int^{\infty} \lambda_1^{g_1} \dots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2} - t_{l_1}) + \dots + \lambda_k(t_{l_k} - t_{l_{k-1}})\}] dt_{l_k} \dots dt_1 \\
& - \int_{\nu_1}^{t_2} \dots \int_{\nu_{l_k-1}}^{t_{l_k}} \int_{x}^{\infty} \lambda_1^{g_1} \dots \lambda_k^{g_k} \exp[-\{\lambda_1 t_{l_1} + \lambda_2(t_{l_2} - t_{l_1}) + \dots + \lambda_k(t_{l_k} - t_{l_{k-1}})\}] dt_{l_k} \dots dt_1 \\
= & \mathfrak{B}_{l_k-1}\left(h^{-1}(y_k - y); \nu_1, \dots, \nu_{l_k-1}\right) - \mathfrak{B}_{l_k-1}\left(x; \nu_1, \dots, \nu_{l_k-1}\right). \qquad (17)
\end{aligned}
$$

Substituting (16) and (17) into (14), formula (12) directly follows. Hence, the asserted result holds true. □

Corollary B.2 *When $y = 0$, the finite-time ruin probability given by Proposition B.1 coincides with the results of Theorem 2.1 in Ignatov and Kaishev (2012b).*

Proof: Following an analogous derivation as in Appendix A and noting $l_{j(k)} \leq k < l_{j(k)+1}$, where $l_{j(k)} = g_1 + \dots + g_{j(k)}$, the result directly follows. □